

FIGURE 1

GGGGCTTCGGCGCCAGCGGCCAGCGCTAGTCGGTCTGGTAAGGATTTACAAAAGGTGCAGGTATG
AGCAGGTCTGAAGACTAACATTTTGTGAAGTTGTAACAGAAAACCTGTTAGAAATGTTGGTGGT
TTCAGCAAGGCCTCAGTTTCCTTCCTTCAGCCCTTGTAATTTGGACATCTGCTGCTTTCATATTT
TCATACATTACTGCAGTAACACTCCACCATATAGACCCGGCTTTACCTTATATCAGTGACACTGG
TACAGTAGCTCCAGAAAAATGCTTATTTGGGGCAATGCTAAATATTGCGGCAGTTTTATGCATTG
CTACCATTTATGTTTCGTTATAAGCAAGTTCATGCTCTGAGTCCTGAAGAGAACGTTATCATCAA
TTAAACAAGGCTGGCCTTGTACTTGAATACTGAGTTGTTTAGGACTTTCTATTGTGGCAAACCTT
CCAGAAAAACAACCTTTTGTGTCACATGTAAGTGGAGCTGTGCTTACCTTTGGTATGGGCTCAT
TATATATGTTTGTTCAGACCATCCTTTCTACCAAATGCAGCCCAAAATCCATGGCAAACAAGTC
TTCTGGATCAGACTGTTGTTGGTTATCTGGTGTGGAGTAAGTGCACCTTAGCATGCTGACTTGCTC
ATCAGTTTTGCACAGTGGCAATTTTGGGACTGATTTAGAACAGAACTCCATTGGAACCCCGAGG
ACAAAGGTTATGTGCTTCACATGATCACTACTGCAGCAGAATGGTCTATGTCATTTTCCTTCTTT
GGTTTTTTCCTGACTTACATTCGTGATTTTCAGAAAATTTCTTTACGGGTGGAAGCCAATTTACA
TGGATTAACCTCTATGACACTGCACCTTGCCCTATTAACAATGAACGAACACGGCTACTTTCCA
GAGATATTTGATGAAAGGATAAAATATTTCTGTAATGATTATGATTCTCAGGGATTGGGGAAAGG
TTCACAGAAGTTGCTTATTCTTCTCTGAAATTTTCAACCACTTAATCAAGGCTGACAGTAACACT
GATGAATGCTGATAATCAGGAAACATGAAAGAAGCCATTTGATAGATTATTCTAAAGGATATCAT
CAAGAAGACTATTAAAAACACCTATGCCTATACTTTTTTATCTCAGAAAATAAAGTCAAAGACT
ATG

FIGURE 2

<subunit 1 of 1, 266 aa, 1 stop

<MW: 29766, pI: 8.39, NX(S/T): 0

MWFFQQGLSFLPSALVIWTSAAFIIFSITAVTLHHIDPALPYISDTGTVAPEKCLFGAMLNIAAV
LCIATIIYVRYKQVHALSPEENVIIKLNKAGLVLGILSCLGLSIVANFQKTTLFAAHVSGAVLTFG
MGSLYMFVQTILSYQMOPKIHGKQVFWIRLLLVICWGSALSMLTCSSVLHSGNFGTDLEQKLHW
NPEDKGYVLHMITTAAEWSMSFSFFGFLTYIRDFQKISLRVEANLHGLTLYDTAPCPINNERTR
LLSRDI

Important features:

Type II transmembrane domain:

amino acids 13-33

Other Transmembrane domains:

amino acids 54-73, 94-113, 160-180, 122-141

N-myristoylation sites.

amino acids 57-63, 95-101, 99-105, 124-130, 183-189

FIGURE 3

CGGACGCGTGGGCGGACGCGTGGGGGAGAGCCGAGTCCCGGCTGCAGCACCTGGGAGAAGGCAGACC
GTGTGAGGGGGCCTGTGGCCCCAGCGTGCTGTGGCCTCGGGGAGTGGGAAGTGGAGGCAGGAGCCTTC
CTTACACTTCGCCATGAGTTTCCTCATCGACTCCAGCATCATGATTACCTCCCAGATACTATTTTTTG
GATTTGGGTGGCTTTTCTTCATGCGCCAATTGTTTAAAGACTATGAGATACGTGATATGTTGTACAG
GTGATCTTCTCCGTGACGTTTGCATTTTCTTGACCATTGTTGAGCTCATCATCTTTGAAATCTTAGG
AGTATTGAATAGCAGCTCCCGTTATTTTCACTGGAAAATGAACCTGTGTGTAATTCTGCTGATCCTGG
TTTTCATGGTGCCTTTTTACATTGGCTATTTTATTGTGAGCAATATCCGACTACTGCATAAACACGA
CTGCTTTTTTCTGTCTCTTATGGCTGACCTTTATGTATTTCTTCTGAAACTAGGAGATCCCTTTCC
CATTCTCAGCCCAAACATGGGATCTTATCCATAGAACAGCTCATCAGCCGGGTGGTGTGATTGGAG
TGACTCTCATGGCTCTTCTTTCTGGATTGGTGTGCTCAACTGCCATACACTTACATGTCTTACTTC
CTCAGGAATGTGACTGACACGGATATTCTAGCCCTGGAACGGCGACTGCTGCAAACCATGGATATGAT
CATAAGCAAAAAGAAAAGGATGGCAATGGCAGGAGAACAAATGTCCAGAAGGGGGAAGTGCATAACA
AACCATCAGTTTCTGGGAATGATAAAAAGTGTACCCTTCAGCATCAGGAAGTAAAATCTTACT
CTTATTCAACAGGAAGTGGATGCTTTGGAAGAATTAAGCAGGCAGCTTTTTCTGGAAACAGCTGATCT
ATATGCTACCAAGGAGAGAATAGAATACTCCAAAACCTCAAGGGGAAATATTTTAATTTTCTTGGTT
ACTTTTTCTCTATTTACTGTGTTTGGAAAATTTTCATGGCTACCATCAATATTGTTTTTGATCGAGTT
GGGAAAACGGATCCTGTCAAGAGGCATTGAGATCACTGTGAATTATCTGGGAATCCAATTTGATGT
GAAGTTTGGTCCCAACACATTTCTTCATTCTTGTGGAATAATCATCGTCACATCCATCAGAGGAT
TGCTGATCACTCTTACCAAGTTCTTTTATGCCATCTCTAGCAGTAAGTCTCCAATGTCATTGTCTGT
CTATTAGCACAGATAATGGGCATGTACTTTGTCTCTCTGTGCTGCTGATCCGAATGAGTATGCCTTT
AGAATACCGCACCATAATCACTGAAGTCTTGGAGAACTGCAGTTCAACTTCTATCACCGTTGGTTTG
ATGTGATCTTCTGGTCAAGCGCTCTCTCTAGCATACTCTTCTCTATTTGGCTCACAACAGGCACCA
GAGAAGCAAATGGCACCTTGAACTTAAGCCTACTACAGACTGTTAGAGGCCAGTGGTTTCAAATTTA
GATATAAGAGGGGGGAAAAATGGAACCAGGCCTGACATTTTATAAACAAACAAAATGCTATGGTAGC
ATTTTTCACCTTCATAGCATACTCCTTCCCGTCAGGTGATACTATGACCATGAGTAGCATCAGCCAG
AACATGAGAGGGGAGAACTAACTCAAGACAATACTCAGCAGAGAGCATCCCGTGTGGATATGAGGCTGG
TGTAGAGCGGAGAGGAGCCAAGAACTAAAGGTGAAAAATACACTGGAACCTCTGGGGCAAGACATGT
CTATGGTAGCTGAGCCAAACACGTAGGATTTCCGTTTTAAGGTTACATGGAAAAGGTTATAGCTTTG
CCTTGAGATTGACTCATTAAATCAGAGACTGTAACAAAAAAAAAAAAAAAAAAAAAGGGCGGCCGCG
ACTCTAGAGTCGACCTGCAGAAGCTTGCCGCCATGGCCCACTTGTTTATTGCAGCTTATAATG

FIGURE 4

MSFLIDSSIMITSQILFFFGFWLFFMRQLFKDYEIRQYVVQVIFSVTFAFSCTMFELIIFEILGV
LNSSSRYPFHWMNLCVILLILVFMVPFYIGYFIVSNIRLLHKQRLLFSCLLWLTFMYFFWKLGD
FPILSPKHGILSIEQLISRVGVIQVTLALLSGFGAVNCPYTYMSYFLRNVTDTDILALERRLLQ
TMDMIISKKKRMAMARRTMFQKGEVHNKPSGFWGMIKSVTTSASGSENLTLIQQEVDALIELSRQ
LFLETADLYATKERIEYSKTFKGKYFNFLGYFFSIYCVWKIFMATINIVFDRVGKTDVTRGIEI
TVNYLGIQFDVKFWSQHISFILVGIIIVTSIRGLLITLTKFFYAISSSKSSNVIVLLAQIMGY
FVSSVLLIRMSMPLEYRTIITEVLGELQFNFYHRWFDVIFLVLSALSSILFLYLAHQAPEKQMAP

Important features:

Signal peptide:

amino acids 1-23

Potential transmembrane domains:

amino acids 37-55, 81-102, 150-168, 288-311, 338-356, 375-398,
425-444

N-glycosylation sites.

amino acids 67-70, 180-183 and 243-246

Eukaryotic cobalamin-binding proteins

amino acids 151-160

FIGURE 5

AGCAGGAAATCCGGATGTCTCGTTATGAAGTGGAGCAGTGAGTGTGAGCCTCAACATAGTTCC
AGAAGTCTCCATCCGGACTAGTTATTGAGCATCTGCCTCTCATATCACCAGTGGCCATCTGAGGT
GTTTCCCTGGCTCTGAAGGGTAGGCACGATGGCCAGGTGCTTCAGCCTGGTGTGCTTCTCACT
TCCATCTGGACCACGAGGCTCCTGGTCCAAGGCTCTTTGCGTGCAGAAGAGCTTTCCATCCAGGT
GTCATGCAGAATTATGGGGATCACCCCTTGTGAGCAAAAAGGCGAACCAGCAGCTGAATTCACAG
AAGCTAAGGAGGCCCTGTAGGCTGCTGGGACTAAGTTTGGCCGGCAAGGACCAAGTTGAAACAGCC
TTGAAAGCTAGCTTTGAAACTTGACGCTATGGCTGGGTGGAGATGGATTTCGTGGTCACTCTAG
GATTAGCCCAAAACCCCAAGTGTGGGAAAAATGGGGTGGGTGCTCTGATTTGGAAGGTTCCAGTGA
GCCGACAGTTTGCAGCCTATTGTTACAACCTCATCTGATACTTGGACTAAGTTCGTGCATTCCAGAA
ATTATCACCACCAAGATCCCATTCAACACTCAAAGTGCACACAAACAACAGAATTTATTGT
CAGTGACAGTACCTACTCGGTGGCATCCCCCTTACTCTACAATACCTGCCCTACTACTACTCCTC
CTGCTCCAGCTTCCACTTCTATTCCACGGAGAAAAAATTGATTTGTGTACAGAAAGTTTATG
GAACTAGCACCATGTCTACAGAACTGAACCATTTGTTGAAAAATAAGCAGCATTCAAGAATGA
AGCTTGGGTTTGGAGGTGTCCCACGGCTCTGCTAGTGTCTGCTCTCCTCTTCTTTGGTGTG
CAGCTGCTCTGGATTGTTGCTATGTCAAAGGTATGTGAAGGCTTCCCTTTTACAAACAAGAAT
CAGCAGAAGGAAATGATCGAAACCAAGTAGTAAAGGAGGAGAAGGCCAATGATAGCAACCCTA
TGAGGAATCAAAGAAACTGATAAAACCCAGAAGAGTCCAAGAGTCCAAGCAAACTACCGTGC
GATGCCTGGAAGCTGAAGTTTAGATGAGACAGAAATGAGGAGACACACCTGAGGCTGGTTCTTT
CATGCTCCTTACCCTGCCCCAGCTGGGGAAATCAAAAGGGCCAAAGAACCAGAAAGAAAGTCCA
CCCTTGGTTCCTAAGTGAATCAGCTCAGGACTGCCATTGGACTATGGAGTGCACCAAGAGAAT
GCCCTTCTCCTTATTGTAACCCCTGTCTGGATCCTATCCTCCTACCTCCAAAGCTTCCCACGGCCT
TTCTAGCCTGGCTATGTCTAATAATATCCCACTGGGAGAAAGGAGTTTGCAGAGTGAAGGAC
CTAAACATCTCATCAGTATCCAGTGGTAAAAAGGCTCCTGGCTGTCTGAGGCTAGGTGGGTTG
AAAGCCAGGAGTCACTGAGACCAAGGCTTTCTCTACTGATTCCGCAGCTCAGACCTTTCTTCA
GCTCTGAAAGAGAAACACGTATCCACCTGACATGTCTTCTGAGCCCGGTAAGAGCAAAAGAAT
GGCAGAAAAGTTTAGCCCTGAAAGCCATGGAGATTCTCATAACTTGAGACCTAATCTCTGTAA
GCTAAATAAAGAAATAGAACAAGGCTGAGGATACGACAGTACACTGTGAGCAGGACTGTAAAC
ACAGACAGGGTCAAAGTGTTTTCTCTGAACACATTGAGTTGGAATCACTGTTTGAACACACACA
CTTACTTTTTCTGGTCTCTACCACTGCTGATATTTCTCTAGGAAATATACTTTTACAAGTAACA
AAAATAAAACTCTTATAAATTTCTATTTTATCTGAGTTACAGAAATGATTACTAAGGAAGATT
ACTCAGTAATTTGTTTAAAAAGTAATAAAATTCACAAACATTTGCTGAATAGCTACTATATGTC
AAGTGTGTGCAAGGTATTACACTCTGTAATTGAATATTATTCCTCAAAAAATTGCACATAGTAG
AACGCTATCTGGGAAGCTATTTTTTTCAGTTTTGATATTTCTAGCTTATCTACTTCCAACTAAT
TTTTATTTTTGCTGAGACTAATCTTATTCATTTTCTCTAATATGGCAACCATTATAACCTTAAT
TATTATTAACATACCTAAGAAGTACATTGTTACCTCTATATACCAAGCACATTTTAAAGTGCC
ATTAACAAATGTATCACTAGCCCTCCTTTTTTCCAACAAGAAGGACTGAGAGATGCAGAAATATT
TGTGACAAAAATTAAGCATTTAGAAACTT

FIGURE 6

MARCFSLVLLLSIWTTTRLIVQGSRLAEELSIQVSCRIMGITLVSKKANQQLNFTAEKEACRLLG
LSLAGKDQVETALKASFETCSYGWVGDFVVISRISPNPKCGKNGVGVLWKVPVSRQFAAYCYN
SSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPAPTTTPAPASTSIPR
RKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTALLVLALLFFGAAAGLGFCYVK
RYVKAFPFNTKNQKEMIETKVVKEEKANDSNPNESKKTDKNPESKSPSKTTVRCLEAEV

Signal sequence:

amino acids 1-16

Transmembrane domain:

amino acids 235-254

N-glycosylation site.

amino acids 53-57, 130-134, 289-293

Casein kinase II phosphorylation site.

amino acids 145-149, 214-218

Tyrosine kinase phosphorylation site.

amino acids 79-88

N-myristoylation site.

amino acids 23-29, 65-71, 234-240, 235-239, 249-255, 253-259

FIGURE 7

CGCCGCGCTCCCGCACCCGCGGCCCGCCACC GCGCGCTCCCGCATCTGCACCCGAGCCCGG
GGCTCCCGGGCGGAGCGAGCAGATCCAGTCCGGCCCGCAGCGCAACTCGGTCCAGTCGGGGCGG
CGGTGCGGGCGCAGAGCGGAGATCGAGCGGCTTGGGGCCACCCTGCTGTGCCTGCTGCTGGCGG
CGGCGGTCCCCACGGCCCCCGCGCCGCTCCGACGGCGACCTCGGCTCCAGTCAAGCCCGGCCCG
GCTCTCAGCTACCCGCGAGGAGGAGGCCACCCTCAATGAGATGTTCCGCGAGGTTGAGGAACTGAT
GGAGACACGCAGCACAAATTGCGCAGCGCGGTGGAAGAGATGGAGGCAGAAGAAGCTGCTGCTA
AAGCATCATCAGAAGTGAACCTGGCAAACCTTACCTCCCAGCTATCACAATGAGACCAACACAGAC
ACGAAGGTTGGAATAATACCATCCATGTGCACCGAGAAATTCACAAGATAACCAACACCAGAC
TGGCAAAATGGTCTTTTCAGAGACAGTTATCAGATCTGTGGGAGACGAAGAAGGCAGAAGGAGCC
ACGAGTGCATCATCGACGAGGACTGTGGGCCAGCATGTACTGCCAGTTTGCCAGCTTCCAGTAC
ACCTGCCAGCCATGCCGGGGCCAGAGGATGCTCTGCACCCGGGACAGTGAGTGCTGTGGAGACCA
GCTGTGTGTCTGGGGTCACTGCACCAAAATGGCCACCAGGGGCGAGCAATGGGACCATCTGTGACA
ACCAGAGGGACTGCCAGCCGGGGCTGTGCTGTGCCTTCCAGAGAGGCTGCTGTTCCTGTGTGC
ACACCCCTGCCGTGGAGGGCGAGCTTTGCCATGACCCGCCAGCCGGCTTCTGGACCTCATCAC
CTGGGAGCTAGAGCCTGATGGAGCCTTGGACCGATGCCCTTGTGCCAGTGGCCTCCTCTGCCAGC
CCCACAGCCACAGCCTGGTGTATGTGTGCAAGCCGACCTTCGTGGGGAGCCGTGACCAAGATGGG
GAGATCCTGCTGCCAGAGAGGTCCCCGATGAGTATGAAGTTGGCAGCTTCATGGAGGAGGTGCG
CCAGGAGCTGGAGGACCTGGAGAGGAGCCTGACTGAAGAGATGGCGCTGGGGGAGCCTGCGGCTG
CCGCCGCTGCACTGCTGGGAGGGGAAGAGATTAGATCTGGACCAGGCTGTGGGTAGATGTGCAA
TAGAAATAGCTAATTTATTTCCCCAGGTGTGTGCTTTAGGCGTGGGCTGACCAGGCTTCTTCCTA
CATCTCTTCCCAGTAAGTTTCCCTCTGGCTTGACAGCATGAGGTGTTGTGCATTGTTCAGCT
CCCCAGGCTGTTCTCCAGGCTTCACAGTCTGGTGCTTGGGAGAGTCAGGCAGGGTTAACTGCA
GGAGCAGTTTGCCACCCCTGTCCAGATTATTGGCTGCTTTGCCTCTACCAGTTGGCAGACAGCCG
TTTGTCTACATGGCTTTGATAATTGTTTGGAGGGAGGAGATGGAAACAATGTGGAGTCTCCCTC
TGATTGGTTTTGGGGAATGTGGAGAAGAGTGCCCTGCTTTGCAAACATCAACCTGGCAAAAATG
CAACAAATGAATTTCCACGCGAGTTCTTTCCATGGGCATAGGTAAGCTGTGCCTTCAGCTGTTGC
AGATGAAATGTTCTGTTACCCCTGCATTACATGTGTTTATTCATCCAGCAGTGTGCTCAGCTCC
TACCTCTGTGCCAGGGCAGCATTTCATATCCAAGATCAATTCCCTCTCTCAGCACAGCCTGGGG
AGGGGGTCATTGTTCTCCTCGTCCATCAGGGATCTCAGAGGCTCAGAGACTGCAAGCTGCTTGCC
CAAGTCACACAGCTAGTGAAGACCAGAGCAGTTTCATCTGTTGTGACTCTAAGCTCAGTGCTCT
CTCCACTACCCACACCAGCCTTGGTGCCACCAAAAGTGCTCCCCAAAAGGAAGGAGAATGGGAT
TTTTCTTGAGGCATGCACATCTGGAATTAAGGTCAAACCTAATTCTCACATCCCTCTAAAAGTAAA
CTACTGTTAGGAACAGCAGTGTCTCACAGTGTGGGGCAGCCGTCCTTCTAATGAAGACAATGAT
ATTGACACTGTCCCTCTTTGGCAGTTGCATTAGTAACCTTGAAAGGTATATGACTGAGCGTAGCA
TACAGGTTAACCTGCAGAAACAGTACTTAGGTAATTGTAGGGCGAGGATTATAAATGAAATTTGC
AAAATCACTTAGCAGCAACTGAAGACAATTATCAACCACGTGGAGAAAATCAAACCGAGCAGGGC
TGTGTGAAACATGGTTGTAATATGCGACTGCGAACACTGAACTCTACGCCACTCCACAATGATG
TTTTCAGGTGTGACTGTTGCCACCATGTATTCATCCAGAGTCTTAAAGTTTAAAGTTGCA
CATGATTGTATAAGCATGCTTTCTTTGAGTTTAAATTATGTATAAACATAAGTTGCATTAGAA
ATCAAGCATAAATCACTTCAACTGCAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 8

MQRLGATLLCLLLAAAVPTAPAPAPTATSAPVKPGPALSYPQEEATLNEMFREVEELMEDTQHKL
RSAVEEMEAEEAAAKASSEVNLANLPPSYHNETNTDTKVGNNTIHVHREIHKITNNQTGQMVFSE
TVITSVGDEEGRRSHECIIDEDCGPSMYCQFASFQYTCQPCRGQRM LCTR DSECCGDQLCVWGHC
TKMATRGSNGTICDNQRDCQPLCCAFQ RGLLEFPVCTPLPVEGELCHDPASRLLDLITWELEPDG
ALDRCPCASGLLCQPHSHSLVYVCKPTFVGSRDQDGEILLPREVPDEYEVGSFMEEVRQ ELEDLE
RSLTEEMALGEPAAAAAALLGGEI

Signal sequence:

amino acids 1-19

N-glycosylation site.

amino acids 96-100, 106-110, 121-125, 204-208

Casein kinase II phosphorylation site.

amino acids 46-50, 67-71, 98-102, 135-139, 206-210, 312-316,
327-331

N-myristoylation site.

amino acids 202-208, 217-223

Amidation site.

amino acids 140-144

FIGURE 9

CGGACGCGTGGGCGGACGCGTGGGGGCTGTGAGAAAGTGCCAATAAATACATCATGCAACCCAC
GGCCCACCTTGTGAACCTCTCGTGCCAGGGCTGATGTGCGTCTTCCAGGGCTACTCATCCAAAG
GCCTAATCCAACGTTCTGTCTTCAATCTGCAAATCTATGGGGTCTGGGGCTCTTCTGGACCCTT
AACTGGGTACTGGCCCTGGGCAATGCGTCCCTCGCTGGAGCCTTGCCTCCTTCTACTGGGCCTT
CCACAAGCCCCAGGACATCCCTACCTTCCCCTTAATCTCTGCCTTCATCCGCACACTCCGTTACC
ACACTGGGTCAATTGGCATTTGGAGCCCTCATCTGACCCTTGTGCAGATAGCCCGGGTCATCTTG
GAGTATATTGACCACAAGCTCAGAGGAGTGCAGAACCCCTGTAGCCCGCTGCATCATGTGCTGTTT
CAAGTGCTGCCTCTGGTGTCTGGAAAAATTTATCAAGTTCCTAAACCGCAATGCATACATCATGA
TCGCCATCTACGGGAAGAATTTCTGTGTCTCAGCCAAAAATGCGTTCATGCTACTCATGCGAAAC
ATTGTCAGGGTGGTCTGCTCTGGACAAAGTACAGACCTGCTGCTGTTCTTTGGGAAGCTGCTGGT
GGTCGGAGGCGTGGGGGCTCTGTCTTCTTTTTTTCTCCGGTTCGCATCCCGGGGCTGGGTAAAG
ACTTTAAGAGCCCCACCTCAACTATTACTGGCTGCCATCATGACCTCCATCCTGGGGGCCTAT
GTCATCGCCAGCGGCTTCTTCAGCGTTTTTCGGCATGTGTGTTGGACACGCTCTTCTCTGCTTCCT
GGAAGACCTGGAGCGGAACAACGGCTCCCTGGACCGGCCCTACTACATGTCCAAGAGCCTTCTAA
AGATTCTGGGCAAGAAGAACGAGGCGCCCCCGGACAACAAGAAGAGGAAGAAGTGACAGCTCCGG
CCCTGATCCAGGACTGCACCCACCCCCACCGTCCAGCCATCCAACCTCACTTCGCCTTACAGGT
CTCCATTTTGTGGTAAAAAAGGTTTTAGGCCAGGCGCCGTGGCTCACGCCTGTAATCCAACACT
TTGAGAGGCTGAGGCGGGCGGATCACCTGAGTCAGGAGTTCGAGACCAGCCTGGCCAACATGGTG
AAACCTCCGTCTCTATTAAAAATACAAAAATTAGCCGAGAGTGGTGGCATGCACCTGTATCCCA
GCTACTCGGGAGGCTGAGGCAGGAGAATCGCTTGAACCCGGGAGGCAGAGGTTGCAGTGAGCCGA
GATCGCGCCACTGCACTCCAACCTGGGTGACAGACTCTGTCTCCAAAACAAAACAAAACAAA
AAGATTTTATTAAAGATATTTTGTAACTC

FIGURE 10

RTRGRTRGGCEKVPINTSCNPTAHLVNSSCPGLMCVFQGYSSKGLIQRSVFNLQIYGVLGFWTL
NWVLALGQCVLGAFASFYWAFHKPQDIPTFPLISAFIRTLRYHTGSLAFGALILTLVQIARVIL
EYIDHKLRGVQNPVARCIMCCFKCLWCLEKFIKFLNRNAYIMIAIYGKNFCVSAKNAFMLLMRN
IVRVVLDKVTDLLLFFGKLLVGGVGVLSEFFFSGRIPGLGKDFKSPHLNYYWLPIMTSILGAY
VIASGFFSVFGMCDTLFLCFLEDLERNNGSLDRPYMSKSLKILGKKNEAPPDNKKRKK

Important features:

Transmembrane domains:

amino acids 57-80 (type II), 110-126, 215-231, 254-274

N-glycosylation sites.

amino acids 16-20, 27-31, 289-293

Hypothetical YBR002c family proteins.

amino acids 276-288

Ammonium transporters proteins.

amino acids 204-231

N-myristoylation sites.

amino acids 60-66, 78-84

Amidation site.

amino acids 306-310

FIGURE 11

GCCCCGCGCCCGGCGCCGGGCGCCCGAAGCCGGGAGCCACCGCCCATGGGGGCCTGCCTGGGAGCCTGC
TCCCTGCTCAGCTGCGCGTCCTGCCTCTGCGGCTCTGCCCCCTGCATCCTGTGCAGCTGCTGCCCCGC
CAGCCGCAACTCCACCGTGAGCCGCCTCATCTTCACGTTCTTCCTCTTCTGGGGGTGCTGGTGTCCA
TCATTATGCTGAGCCCGGGCGTGGAGAGTCAGCTCTACAAGCTGCCCTGGGTGTGTGAGGAGGGGGCC
GGGATCCCCACCGTCCTGCAGGGCCACATCGACTGTGGCTCCCTGCTTGGCTACCGCGCTGTCTACCG
CATGTGCTTCGCCACGGCGGCCTTCTTCTTCTTCTTTTACCCCTGCTCATGCTCTGCGTGAGCAGCA
GCGGGGACCCCGGGCTGCCATCCAGAATGGGTTTTGGTTCTTTAAGTTCCTGATCCTGGTGGGCCTC
ACCGTGGGTGCCTTCTACATCCCTGACGGCTCTTACCAACATCTGGTTCTACTTCGGCGTCGTGGG
CTCCTTCTCTTCATCCTCATCCAGCTGGTGTGCTCATCGACTTTGCGCACTCCTGGAACCAGCGGT
GGCTGGGCAAGCCGAGGAGTGCGATTCCCGTGCCTGGTACGCAGGCCTCTTCTTCTTCACTCTCCTC
TTTACTTGCTGTGATCGCGGCGCTGGCGCTGATGTTTATGTACTACACTGAGCCCAGCGGCTGCCA
CGAGGGCAAGGTCTTCATCAGCCTCAACCTCACCTTCTGTGTCTGCGTGTCCATCGCTGTCTCTGC
CCAAGTCCAGGACGCCCAGCCCAACTCGGGTCTGCTGCAGGCCTCGGTATCACCCCTCTACACCATG
TTTGTACCTGGTCAGCCCTATCCAGTATCCCTGAACAGAAATGCAACCCCATTTGCCAACCAGCT
GGGCAACGAGACAGTTGTGGCAGGCCCCGAGGGCTATGAGACCCAGTGGTGGGATGCCCCGAGCATTG
TGGGCCTCATCATCTTCTCCTGTGCACCCCTTTCATCAGTCTGCGCTCCTCAGACCACCGGCAGGTG
AACAGCCTGATGCAGACCGAGGAGTGCCACCTATGCTAGACGCCACACAGCAGCAGCAGCAGGT
GGCAGCCTGTGAGGGCCGGGCCTTTGACAACGAGCAGGACGGCGTCACCTACAGCTACTCCTTCTTCC
ACTTCTGCCTGGTGTGCGCTCACTGCACGTCATGATGACGCTACCAACTGGTACAAGCCCGGTGAG
ACCCGGAAGATGATCAGCACGTGGACCGCCGTGTGGGTGAAGATCTGTGCCAGCTGGGCAGGGCTGCT
CCTCTACCTGTGGACCCCTGGTAGCCCCACTCCTCCTGCGCAACCGCGACTTCAGCTTGAGGCAGCCTCA
CAGCCTGCCATCTGGTGCCTCCTGCCACCTGGTGCCTCTCGGCTCGGTGACAGCCAACCTGCCCCCTC
CCCACACCAATCAGCCAGGCTGAGCCCCACCCCTGCCCCAGCTCCAGGACCTGCCCCTGAGCCGGGC
CTTCTAGTCGTAGTGCCCTCAGGGTCCGAGGAGCATCAGGCTCCTGCAGAGCCCCATCCCCCGCCAC
ACCCACACGGTGGAGCTGCCTCTTCCCTTCCCTCCTCCTGTTGCCATACTCAGCATCTCGGATGAA
AGGGCTCCCTTGTCTCAGGCTCCACGGGAGCGGGGTGCTGGAGAGAGCGGGGAAGTCCCACCACAG
TGGGGCATCCGGCACTGAAGCCCTGGTGTCTGTCACGTCCCCAGGGGACCTGCCCCCTTCTCTG
GACTTCGTGCCTTACTGAGTCTCTAAGACTTTTTCTAATAACAAGCCAGTGCCTGTAAAAAAA

FIGURE 12

MGACLGACSLSCASCSCGAPCILSCCPCASRNSTVSRIFTFFFLGVLVSIIMLSPGVESQL
YKLPWVCEEGAGIPTVLQGHIDCGSLLGYRAVYRMCFATAAFFFFFFFTLLMLCVSSSRDPRAAIQ
NGFWFFKFLILVGLTVGAFYIPDGSFTNIWFYFGVVGSLFILIQLVLLIDFAHSWNQRWLKAE
ECDSRAWYAGLFFFTLLFYLLSIAAVALMFMYYTEPSGCHEGKVFISLNLTCVCVSIAAVLPKV
QDAQPNSGLLQASVITLYTMFVTWSALSSIPEQKCNPHLPTQLGNETVVAGPEGYETQWWDAPSI
VGLIIFLLCTLFISLRSSDHRQVNSLMQTEECPPMLDATQQQQQVAACEGRAFDNEQDGVITYSY
SFFHFCLVLASLHVMMTLTNWYKPGETRKMISTWTAVWVKICASWAGLLLYLWTLVAPLLLRNRD
FS

Signal sequence:

amino acids 1-20

Transmembrane domains:

amino acids 40-58, 101-116, 134-150, 162-178, 206-223, 240-257,
272-283, 324-340, 391-406, 428-444

FIGURE 13

CGGGCCAGCCTGGGGCGGCCGGCCAGGAACCAACCCGTTAAGGTGTCTTCTCTTTAGGGATGGTGA
GGTTGGAAAAAGACTCCTGTAACCCCTCCTCCAGGATGAACCACCTGCCAGAAGACATGGAGAACG
CTCTCACCGGGAGCCAGAGCTCCCATGCTTCTCTGCGCAATATCCATTCCATCAACCCACACAA
CTCATGGCCAGGATTGAGTCCATGAAGGAAGGGAAAAGAAAGGCATATCTGATGTCAGGAGGAC
TTTCTGTTTGTGTCACCTTTGACCTCTTATTCGTAACATTACTGTGGATAATAGAGTTAAATG
TGAATGGAGGCATTGAGAACACATTAGAGAAGGAGGTGATGCAGTATGACTACTATTCTTCATAT
TTTGATATATTTCTTCTGGCAGTTTTTCGATTTAAAGTGTTAACTTGCATATGCTGTGTGCAG
ACTGCCCATTTGGTGGGCAATAGCGTTGACAACGGCAGTGACCAGTGCCTTTTTACTAGCAAAAG
TGATCCTTTTGAAGCTTTTCTCTCAAGGGGCTTTTGGCTATGTGCTGCCCATCATTTCATTCATC
CTTGCTGGATTGAGACGTGGTTCCTGGATTTCAAAGTGTTACCTCAAGAAGCAGAAGAAGAAAA
CAGACTCCTGATAGTTCAGGATGCTTCAGAGAGGGCAGCACTTATACCTGGTGGTCTTTCTGATG
GTCAGTTTTATTCCCCTCCTGAATCCGAAGCAGGATCTGAAGAAGCTGAAGAAAAACAGGACAGT
GAGAAACCACTTTTGAAGTATGAGTACTACTTTTGTTAAATGTGAAAAACCCCTCACAGAAAGTC
ATCGAGGCAAAAAGAGGCAGGCAGTGAGTCTCCCTGTCGACAGTAAAGTTGAAATGGTGACGTC
CACTGCTGGCTTTTATGAACAGCTAATAAAGATTTATTTATTTGTAATACCTCACAAACGTTGTAC
CATATCCATGCACATTTAGTTGCCTGCCTGTGGCTGGTAAGGTAATGTCATGATTCATCCTCTCT
TCAGTGAGACTGAGCCTGATGTGTTAACAAATAGGTGAAGAAAGTCTTGTGCTGTATTCTTAATC
AAAAGACTTAATATATTGAAGTAACACTTTTTTAGTAAGCAAGATACCTTTTTATTTCATTCAC
AGAATGGAATTTTTTGTTCATGTCTCAGATTTATTTGTATTTCTTTTTAACACTCTACATT
TCCCTTGTTTTTTAACTCATGCACATGTGCTCTTTGTACAGTTTTTAAAAAGTGAATAAAATCTG
ACATGTCAATGTGGCTAGTTTTATTTTTCTTGTTTTGCATTATGTGTATGCCCTGAAGTGTGGA
CTTGCAAAAGGGGAAGAAAGGAATTGCGAATACATGTAAATGTCACCAGACATTTGTATTATTT
TTATCATGAAATCATGTTTTCTCTGATTGTCTGAAATGTTCTAAATACTCTTATTTGAATGC
ACAAAATGACTTAAACCATTATATCATGTTTCCTTTGCGTTCAGCCAATTTCAATTAAAATGAA
CTAAATTAAAAA

FIGURE 14

MNHLPEDMENALTGSQSSHASLRNIHSINPTQLMARIESYEGREKKGISDVRRTFCLEFVTFDLLF
VTLLWIIELNVNGGIENTLEKEVMQYDYSSYFDIFLLAVFRFKVLILAYAVCRLRHWWAIALTT
AVTSAFLLAKVILSKLFSQGAFGYVLPPIISFILAWIETWFDFKVLQPQEAEEENRLLIVQDASER
AALIPGGLSDGQFYSPPESEAGSEEAEKQDSEKPLLEL

Important features of the protein:

Signal peptide:

amino acids 1-20

Transmembrane domains:

amino acids 54-72, 100-118, 130-144, 146-166

N-myristoylation sites.

amino acids 14-20, 78-84, 79-85, 202-208, 217-223

FIGURE 15

ACTCGAACGCAGTTGCTTCGGGACCCAGGACCCCTCGGGCCCGACCCGCCAGGAAAGACTGAGG
 CCGCGGCCTGCCCCGCCCCGGCTCCCTGCGCCGCGCGCCCTCCCGGGACAGAAGATGTGCTCCAG
 GGTCCCTCTGCTGCTGCCGCTGCTCCTGCTACTGGCCCTGGGGCCTGGGGTGCAGGGCTGCCAT
 CCGGCTGCCAGTGACGCCAGCCACAGACAGTCTTCTGCACTGCCCCGCCAGGGGACCACGGTGCCC
 CGAGACGTGCCACCCGACACGGTGGGGCTGTACGTCTTTGAGAACGGCATCACCATGCTCGACGC
 AGGCAGCTTTGCCGGCCTGCCGGGCTGCAGTCTCTGGACCTGTACAGAACCAGATCGCCAGCC
 TGCCACGCGGGGTCTTCCAGCCACTCGCCAACCTCAGCAACCTGGACCTGACGGCCAACAGGCTG
 CATGAAATCACCAATGAGACCTTCCGTGGCCTGCGGCGCCTCGAGCGCCTCTACCTGGGCAAGAA
 CCGCATCCGCCACATCCAGCCTGGTGCCTTCGACACGCTCGACCGCCTCCTGGAGCTCAAGCTGC
 AGGACAACGAGCTGCGGGCACTGCCCCCGCTGCGCCTGCCCCGCTGCTGCTGCTGGACCTCAGC
 CACAACAGCCTCCTGGCCCTGGAGCCCGGCATCCTGGACACTGCCAACGTGGAGGCGCTGCGGCT
 GGCTGGTCTGGGGCTGCAGCAGCTGGACGAGGGGCTCTTACGCCCTTGCAGAACCTCCACGACC
 TGGATGTGTCCGACAACAGCTGGAGCGAGTGCCACCTGTGATCCGAGGCTCCGGGGCTGACG
 CGCCTGCGGCTGGCCGGCAACACCCGATTGCCCCAGCTGCGGCCGAGGACCTGGCCGCGCTGGC
 TGCCCTGCAGGAGCTGGATGTGAGCAACCTAAGCCTGCAGGCCCTGCCTGGCGACCTCTCGGGCC
 TCTTCCCCCGCTGCGGCTGCTGGCAGCTGCCCGCAACCCCTTCAACTGCGTGTGCCCCCTGAGC
 TGGTTTGGCCCCCTGGGTGCGCGAGAGCCACGTCACTGCGCAGCCCTGAGGAGACGCGCTGCCA
 CTTCCCCGCCAAGAACGCTGGCCGCTGCTCCTGGAGCTTGACTACGCCGACTTTGGCTGCCAG
 CCACCACCACACAGCCACAGTGCCACACGAGGCGCGTGGTGCGGGAGCCACAGCCTTGTCT
 TCTAGCTTGGCTCCTACCTGGCTTAGCCCCACAGCGCCGGCCACTGAGGCCCCAGCCCGCCCTC
 CACTGCCCCACCGACTGTAGGGCTGTCCCCCAGCCCCAGGACTGCCACCGTCCACCTGCCTCA
 ATGGGGGCACATGCCACCTGGGGACACGGCACCTGGCGTGCTTGTGCCCCGAAGGCTTCACG
 GGCTGTACTGTGAGAGCCAGATGGGGCAGGGGACACGGCCAGCCCTACACCAGTCACGCCGAG
 GCCACCACGGTCCCTGACCCTGGGCATCGAGCCGGTGAGCCCCACCTCCCTGCGCGTGGGGCTGC
 AGCGCTACCTCCAGGGGAGCTCCGTGCAGCTCAGGAGCCTCCGTCTCACCTATCGCAACCTATCG
 GGCCCTGATAAGCGGCTGGTGACGCTGCGACTGCCTGCCTCGCTCGCTGAGTACACGGTACCCCA
 GCTGCGGCCCAACGCCACTTACTCCGTCTGTGTGTCATGCCTTTGGGGCCCGGGCGGGTGCCGGAGG
 GCGAGGAGGCTGCGGGGAGGCCCATACACCCAGCCGTCCACTCCAACCAGCCCCAGTCAACC
 CAGGCCCGGAGGGCAACCTGCCGCTCCTCATGCGCCCGCCCTGGCCGCGGTGCTCCTGGCCGC
 GCTGGCTGCGGTGGGGGCGAGCCTACTGTGTGCGGCGGGGGCGGGCCATGGCAGCAGCGGCTCAGG
 ACAAAGGGCAGGTGGGGCCAGGGGCTGGGCCCTGGAAGTGGAGGGAGTGAAGTCCCTTGGAG
 CCAGGCCCGAAGGCAACAGAGGGCGGTGGAGAGGCCCTGCCAGCGGGTCTGAGTGTGAGGTGCC
 ACTCATGGGCTTCCAGGGCTGGCTCCAGTCAACCCCTCCACGCAAGCCCTACATCTAAGCCA
 GAGAGAGACAGGGCAGCTGGGGCCGGGCTCTCAGCCAGTGAGATGGCCAGCCCCCTCCTGCTGCC
 ACACCACGTAAGTTCTCAGTCCCAACCTCGGGGATGTGTGCAGACAGGGCTGTGTGACCACAGCT
 GGGCCCTGTTCCCTCTGGACCTCGGTCTCCTCATCTGTGAGATGCTGTGGCCAGCTGACGAGCC
 CTAACCTCCCCAGAACCAGTGCCTATGAGGACAGTGTCCGCCCTGCCCTCCGCAACGTGCACTC
 CCTGGGCACGGCGGGGCCCTGCCATGTGCTGGTAACGCATGCCTGGGTCTGCTGGGCTCTCCAC
 TCCAGGCGGACCTGGGGGCCAGTGAAGGAAGCTCCCGAAAGAGCAGAGGGAGAGCGGGTAGGC
 GGCTGTGTGACTCTAGTCTTGGCCCCAGGAAGCGAAGGAACAAAAGAACTGGAAGGAAGATGC
 TTTAGGAACATGTTTTGCTTTTTTAAATATATATATTTATAAGAGATCCTTTCCATTATTCT
 GGAAGATGTTTTCAAACCTCAGAGACAAGGACTTTGGTTTTTGTAAAGACAACGATGATATGAA
 GGCTTTTGTAAAGAAAAATAAAGATGAAGTGTGAAA

FIGURE 16

MCSRVPLLLLLLLLLALGPGVQGCPSGCQCSQPQTVFCTARQGTTVPRDVPPDTVGLYVFENGIT
MLDAGSFAGLPGLQLLDLSQNQIASLPSGVFQPLANLSNLDLTANRLHEITNETFRGLRRLERLY
LGKNRIRHIQPGAFDTLDRLLELKLQDNELRALPPLRLPRLLLLDLSHNSLLALEPGILDTANVE
ALRLAGLGLQQLDEGLFSRLRNLDLDVSDNQLERVPPVIRGLRGLTRLRLAGNTRIAQLRPEDL
AGLAALQELDVSNSLSLQALPGDLSGLFPRLRLAAARNPFNCVCPLSWFGPWVRESHVTLASPEE
TRCHFPPKNAGRLLLELDYADFGCPATTTTATVPPTTRPVVREPTALSSSLAPTWLSPTAPATEAP
SPPSTAPPTVGPVPQPDQCPPSTCLNGGTCHLGRHHLACLCPEGFTGLYCESQMGQGTTPSPTP
VTPRPRSLTLGIEPVSPTSLRVGLQRYLQSSVQLRSLRLTYRNLSGPDKRLVTLRLPASLAEY
TVTQLRPNATYSVCMPLGPRVPEGEEACGEAHTPPAVHSNHAPVTQAREGNLPLLIAPALAAV
LLAALAAVGAAYCVRGRAMAAAAQDKGQVGPAGPLELEGVKVPLEPGPKATEGGGEALPSGSE
CEVPLMGFPGLQSPHAKPYI

Important features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 579-599

EGF-like domain cysteine pattern signature.

amino acids 430-442

Leucine zipper pattern.

amino acids 197-219, 269-291

N-glycosylation sites.

amino acids 101-105, 117-121, 273-277, 500-504, 528-532

Tyrosine kinase phosphorylation sites.

amino acids 124-131, 337-345

N-myristoylation sites.

amino acids 23-29, 27-33, 70-76, 142-148, 187-193, 348-354,
594-600, 640-646

FIGURE 17

GCAGCGGCGAGGCGGCGGTGGTGGCTGAGTCCGTGGTGGCAGAGGCGAAGGCGACAGCTCATGCG
GGTCCGGATAGGGCTGACGCTGCTGTGTGTGCGGTGCTGCTGAGCTTGGCCTCGGCGTCCTCGG
ATGAAGAAGGCAGCCAGGATGAATCCTTAGATTCCAAGACTACTTTGACATCAGATGAGTCAGTA
AAGGACCATACTACTGCAGGCAGAGTAGTTGCTGGTCAAATATTTCTTGATTGAGAAGAATCTGA
ATTAGAATCCTCTATTCAAGAAGAGGAAGACAGCCTCAAGAGCCAAGAGGGGAAAGTGTACAG
AAGATATCAGCTTTCTAGAGTCTCCAAATCCAGAAAACAAGGACTATGAAGAGCCAAAGAAAGTA
CGGAAACCAGCTTTGACCGCCATTGAAGGCACAGCACATGGGGAGCCCTGCCACTTCCCTTTTCT
TTTCCTAGATAAGGAGTATGATGAATGTACATCAGATGGGAGGGAAGATGGCAGACTGTGGTGTG
CTACAACCTATGACTACAAAGCAGATGAAAAGTGGGGCTTTTGTGAACTGAAGAAGAGGCTGCT
AAGAGACGGCAGATGCAGGAAGCAGAAATGATGTATCAAACCTGGAATGAAAATCCTTAATGGAAG
CAATAAGAAAAGCCAAAAAAGAGAAGCATATCGGTATCTCCAAAAGGCAGCAAGCATGAACCATA
CCAAAGCCCTGGAGAGAGTGTATATGCTCTTTTATTTGGTGATTACTTGCCACAGAAATATCCAG
GCAGCGAGAGAGATGTTTGAGAAGCTGACTGAGGAAGGCTCTCCCAAGGGACAGACTGCTCTTGG
CTTTCTGTATGCCTCTGGACTTGGTGTAAATCAAGTCAGGCAAAGGCTCTTGATATATATACAT
TTGGAGCTCTTGGGGCAATCTAATAGCCCACATGGTTTTTGGTAAGTAGACTTTTAGTGGAAGGCT
AATAATATTAACATCAGAAGAATTTGTGGTTTATAGCGGCCACAACCTTTTCAGCTTTCATGATC
CAGATTTGCTTGTATTAAGACCAAATATTCAGTTGAACTTCCTTCAAATTCTTGTTAATGGATAT
AACACATGGAATCTACATGTAAATGAAAGTTGGTGGAGTCCACAATTTTCTTTAAATGATTAG
TTTGGCTGATTGCCCCATAAAAAGAGAGATCTGATAAATGGCTCTTTTAAATTTTCTCTGAGTTG
GAATTGTCAGAATCATTTTTTACATTAGATTATCATAATTTTAAAAATTTTCTTTAGTTTTTCA
AAATTTTGTAATGGTGGCTATAGAAAAACAACATGAAATATTATACAATATTTTGCAACAATGC
CCTAAGAATTGTTAAATTCATGGAGTTATTTGTGCAGAAAGTCCAGAGAGCTCTACTTTCTG
TTTTTTACTTTTCATGATTGGCTGTCTTCCCATTTATTCTGGTCATTTATTGCTAGTGACACTGT
GCCTGCTTCCAGTAGTCTCATTTTCCCTATTTTGCTAATTTGTTACTTTTTCTTTGCTAATTTGG
AAGATTAACCTCATTTTTTAATAAAATTATGTCTAAGATTAAAAAAAAAAAAAAAAAAAAAAAAA
AA

FIGURE 18

MRVRIGLTLLLCVLLSLASASSDEEGSQDESLSKTTLTSDSVKDHTTAGRVVAGQIFLDSEESSEL
ESSIQEEEDSLKSQEGESVTEDISFLESPNPENKDYEEPKKVRKPALTAIEGTAHGEPCHFPFLFLDK
EYDECTSDGREDGRLWCATTYDYKADEKWGFCETEEEEAAKRRQMQAEMMYQTGMKI LNGSNKKSQKR
EAYRYLQKAASMNHTKALERVSYALLFGDYL PQNIQAAREMF EKLTEEGSPKGQTALGFLYASGLGVN
SSQAKALVYYTFGALGGNLI AHMVLVSRL

Important features:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 195-199, 217-221, 272-276

Tyrosine kinase phosphorylation site.

amino acids 220-228

N-myristoylation sites.

amino acids 120-126, 253-259, 268-274, 270-274, 285-291, 289-295

Glycosaminoglycan attachment site.

amino acids 267-271

Microbodies C-terminal targeting signal.

amino acids 299-303

Type II fibronectin collagen-binding domain protein.

amino acids 127-169

Fructose-bisphosphate aldolase class-II protein.

amino acids 101-119

FIGURE 19

AATTCAGATTTTAAAGCCCATTTCTGCAGTGGAAATTCATGAACTAGCAAGAGGACACCATCTTCTT
GTATTATACAAGAAAGGAGTGTACCTATCACACACAGGGGGAAAAATGCTCTTTTGGGTGCTAGG
CCTCCTAATCCTCTGTGGTTTTCTGTGGACTCGTAAAGGAAAATAAGATTGAAGACATCACTG
ATAAGTACATTTTTATCACTGGATGTGACTCGGGCTTTGGAACTTGGCAGCCAGAATTTTGAT
AAAAAGGGATTTTCATGTAATCGCTGCCTGTCTGACTGAATCAGGATCAACAGCTTTAAAGGCAGA
AACCTCAGAGAGACTTCGTACTGTGCTTCTGGATGTGACCGACCCAGAGAATGTCAAGAGGACTG
CCCAGTGGGTGAAGAACCAAGTTGGGGAGAAAGGTCTCTGGGGTCTGATCAATAATGCTGGTGT
CCCGCGTGCTGGCTCCCACTGACTGGCTGACACTAGAGGACTACAGAGAACCTATTGAAGTGAA
CCTGTTTGGACTCATCAGTGTGACACTAAATATGCTTCCTTTGGTCAAGAAAGCTCAAGGGAGAG
TTATTAATGTCTCCAGTGTGGAGGTGCGCTTCAATCGTTGGAGGGGGCTATACTCCATCCAAA
TATGCAGTGAAGGTTTCAATGACAGCTTAAGACGGGACATGAAAGCTTTTGGTGTGCACGTCTC
ATGCATTGAACCAGGATTGTTCAAAACAACTTGGCAGATCCAGTAAAGGTAATTGAAAAAAAC
TCGCCATTTGGGAGCAGCTGTCTCCAGACATCAAACAACATATGGAGAAGGTTACATTGAAAAA
AGTCTAGACAACTGAAAGGCAATAAATCCTATGTGAACATGGACCTCTCTCCGGTGGTAGAGTG
CATGGACCACGCTCTAACAAGTCTCTCCCTAAGACTCATTATGCCGCTGGAAAAGATGCCAAAA
TTTTCTGGATACCTCTGTCTCACATGCCAGCAGCTTTGCAAGACTTTTTATTGTTGAAACAGAAA
GCAGAGCTGGCTAATCCCAAGGCAGTGTGACTCAGCTAACCCAAATGTCTCCTCCAGGCTATGA
AATTGGCCGATTTCAAGAACACATCTCCTTTTCAACCCCATTCCTTATCTGCTCCAACCTGGACT
CATTTAGATCGTGCTTATTTGGATTGCAAAGGGAGTCCCACCATCGCTGGTGGTATCCCAGGGT
CCCTGCTCAAGTTTTCTTTGAAAAGGAGGGCTGGAATGGTACATCACATAGGCAAGTCCTGCCCT
GTATTTAGGCTTTGCCTGCTTGGTGTGATGTAAGGGAAATTCAAAGACTTGCCCATTCAAAATGA
TCTTTACCGTGGCCTGCCCCATGCTTATGGTCCCCAGCATTACAGTAACTTGTGAATGTTAAGT
ATCATCTCTTATCTAAATATTAAAGATAAGTCAACCCAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAA

FIGURE 20

MLFWVLGLLILCGFLWTRKGLKIEDITDKYIFITGCDSGFGNLAARTFDKKGPHVIAACLTESG
STALKAETSERLRTVLLDVTDPENVKRTAQWVKNOVGEKGLWGLINNAGVPGVLAPTDWLTLEDY
REPIEVNLFGLISVTLNMLPLVKKAQGRVINVSSVGGRLAIVGGGYTPSKYAVEGFNDLRRDMK
AFGVHVSCEIEPGLFKNLADPVKVEKKLAIWEQLSPDIKQQYGEGYIEKSLDKLKGKNSYVNMD
LSPVVECMDHALTSLFPKTHYAAGKDAKIFWIPLSHMPAALQDFLLLKQKAELANPKAV

Important features of the protein:

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 136-152

N-glycosylation sites.

amino acids 161-163, 187-190 and 253-256

Glycosaminoglycan attachment site.

amino acids 39-42

N-myristoylation sites.

amino acids 36-41, 42-47, 108-113, 166-171, 198-203 and 207-212

FIGURE 21

CTGAGGCGGCGGTAGCATGGAGGGGGAGAGTACGTCGGCGGTGCTCTCGGGCTTTGTGCTCGGCG
CACTCGCTTTCCAGCACCTCAACACGGACTCGGACACGGAAGGTTTTCTTCTTGGGGAAGTAAAA
GGTGAAGCCAAGAACAGCATTACTGATTCCCAAATGGATGATGTTGAAGTTGTTTATACAATTGA
CATTTCAGAAATATATTCATGCTATCAGCTTTTCTAGCTTTTATAATTCTTCAGGCGAAGTAAATG
AGCAAGCACTGAAGAAAATATTATCAAATGTCAAAAAGAATGTGGTAGGTTGGTACAAATTCGGT
CGTCATTTCAGATCAGATCATGACGTTTAGAGAGAGGCTGCTTCACAAAACTTGCAGGAGCATT
TTCAAACCAAGACCTTGTTTTCTGCTATTAACACCAAGTATAATAACAGAAAGCTGCTCTACTC
ATCGACTGGAACATTCCCTTATATAAACCTCAAAAAGGACTTTTTACAGGGTACCTTTAGTGTT
GCCAATCTGGGCATGCTGAACAACTGGGTTATAAACTGTATCAGGTTCTGTATGTCCACTGG
TTTTAGCCGAGCAGTACAAACACACAGCTCTAAATTTTTGAAGAAGATGGATCCTTAAAGGAGG
TACATAAGATAAATGAAATGTATGCTTCATTACAAGAGGAATTAAGAGTATATGCAAAAAAGTG
GAAGACAGTGAACAAGCAGTAGATAAACTAGTAAAGGATGTAAACAGATTAAAACGAGAAATTGA
GAAAAGGAGAGGAGCACAGATTTCAGGCAGCAAGAGAGAAGAACATCCAAAAGACCCTCAGGAGA
ACATTTTTCTTTGTCTAGGCATTACGGACCTTTTTTCCAAATCTGAATTTCTTCATTTCATGTGTT
ATGTCTTTAAAAAATAGACATGTTTCTAAAAGTAGCTGTAACACCAACCACCTCGATGTAGT
AGACAATCTGACCTTAATGGTAGAACACACTGACATTCTGAAGCTAGTCCAGCTAGTACACCAC
AAATCATTAAGCATAAAGCCTTAGACTTAGATGACAGATGGCAATTCAGAGATCTCGGTTGTTA
GATACACAAGACAAACGATCTAAAGCAAATACTGGTAGTAGTAACCAAGATAAAGCATCCAAAT
GAGCAGCCCAGAAACAGATGAAGAAATTGAAAAGATGAAGGGTTTTGGTGAATATTCACGGTCTC
CTACATTTTGATCCTTTTAACCTTACAAGGAGATTTTTTTATTTGGCTGATGGGTAAAGCCAAAC
ATTTCTATGTTTTTACTATGTTGAGCTACTTGCAGTAAGTTCATTTGTTTTTACTATGTTTCACC
TGTTTGCAAGTAATACACAGATAACTCTTAGTGCATTTACTTCACAAAGTACTTTTTCAAACATCA
GATGCTTTTATTTCCAAACCTTTTTTTCACCTTTCACTAAGTTGTTGAGGGGAAGGCTTACACAG
ACACATTCCTTAGAATTGGAAAAGTGAGACCAGGCACAGTGGCTCACACCTGTAATCCCAGCACT
TAGGGAAGACAAGTCAGGAGGATTGATTGAAGCTAGGAGTTAGAGACCAGCCTGGGCAACGTATT
GAGACCATGTCTATTAAAAAATAAAATGAAAAGCAAGAATAGCCTTATTTTCAAATATGGAAA
GAAATTTATATGAAAATTTATCTGAGTCATTAAATTTCTCCTAAGTGATACTTTTTTAGAAGTA
CATTATGGCTAGAGTTGCCAGATAAAATGCTGGATATCATGCAATAAATTTGCAAAACATCATCT
AAAATTTAAAAA

FIGURE 22

MEGESTSAVLSGFVLGALAFQHLNTSDTEGFLLEVKGEAKNSITDSQMDDVEVVYTIDIQKYI
PCYQLFSFYNSSGEVNEQALKKILSNVKKNVVGWYKFRRHSDQIMTFRERLLHKNLQEHFSNQDL
VFLLLTPSIITESCSTHRLEHSLYKPQKGLFHRVPLVVANLGMSEQLGYKTVSGSCMSTGFSRAV
QTHSSKFFEEEDGSLKEVHKINEMYASLQEELKSICKKVEDSEQAVDKLVKDVNRLKREIEKRRGA
QIQAAAREKNIQKDPQENIFLCQALRTFFPNSEFLHSCVMSLKNRHVSKSSCNYNHHLDVVDNLTL
MVEHTDIPEASPASTPQIIKHKALDLDDRWQFKRSRLDQDKRSKANTGSSNQDKASKMSSPET
DEEIEKMKGFGEYSRSPTF

Important features:

Signal peptide:

amino acids 1-19

N-glycosylation sites.

amino acids 75-79, 322-326

N-myristoylation site.

amino acids 184-154

Growth factor and cytokines receptors family.

amino acids 134-150

FIGURE 23

GGCACAGCCGCGCGGAGGGCAGAGTCAGCCGAGCCGAGTCCAGCCGGACGAGCGGACCAGCGCAGGGCAGCCCAA
GCAGCGCGCAGCGAACGCCCGCGCCGCCACACCCCTCTGCGGTCCCCGCGCGCCTGCCACCCTTCCCTCCTTCCCC
GCGTCCCCCGCTCGCGCGCCAGTCAGCTTGCCGGGTTGCTGCCCCGCGAAACCCGAGGTCACCAGCCCGCGCCTCT
GCTTCCCTGGGCGCGCGCCGCTCCACGCCCTCCTTCTCCCTGGCCGCGCCTGGCACCAGGGGACCGTTGCCTGA
CGGAGGCCCAGCTCTACTTTTCCCCCGGCTCTCTCCGCTGCTCGCCTCTTCCACCAACTCCAACCTCCTTCTCCC
TCCAGTCCACTCGCTAGTCCCCGACTCCGCCAGCCCTCGGCCGCTGCGGTAGCGCCGCTTCCCGTCCGGTCCCAAA
GGTGGGAACCGCTCCGCCCGCGCCCGCACCATGGCACGGTTGCGCTTGCCCCGCGCTTCTGTCACCCCTGGCAGTGCTC
AGCGCCGCGCTGCTGGCTGCCGAGCTCAAGTCGAAAAGTTGCTCGGAAGTGCAGCTCTTTACGTGTCCAAGGCTTC
AACAAGAACGATGCCCCCTCCACGAGATCAACGGTGATCATTTGAAGATCTGTCCCAGGGTTCTACCTGCTGCTCT
CAAGAGATGGAGGAGAAGTACAGCCTGCAAAAGTAAAGATGATTTCAAAAGTGTGGTCAGCGAACAGTGCAATCATTTG
CAAGCTGCTTTGCTTACGTTACAAGAAGTTTATGAATTTCTCAAAGAACTACTTGAAATGCAGAGAAATCCCTG
AATGATATGTTTGTGAAGACATATGGCCATTTATACATGCAAAATCTGAGCTATTTAAAGATCTCTTCGTAGAGTTG
AAACGTTACTACGTGGTGGGAAATGTGAACCTGGAAGAAATGCTAAATGACTTCTGGGCTCGCCTCCTGGAGCGGATG
TTCCGCTGGTGAACCTCCAGTACCACCTTACAGATGAGTATCTGGAATGTGTGAGCAAGTATACGGAGCAGCTGAAG
CCCTTCGGAGATGTCCCTCGCAAATGAAGCTCCAGGTTACTCGTGCTTTTGTAGCAGCCCGTACTTTCGCTCAAGGC
TTAGCGGTTGCGGGAGATGTGCTGAGCAAGGTTCTCCGTGGTAAACCCACAGCCAGTGATCCCATGCCCTGTTGAAG
ATGATCTACTGCTCCCACTGCCGGGCTCTCGTGACTGTGAAGCCATGTTACAACACTGCTCAAACATCATGAGAGGC
TGTTTGCCCAACCAAGGGGATCTCGATTTTGAATGGAACAATTTATAGATGCTATGCTGATGGTGGCAGAGAGGCTA
GAGGGTCCTTTCAACATTGAATCGGTGATGATCCCATCGATGTGAAGATTTCTGATGCTATTATGAACATGCAGGAT
AATAGTGTTCAGTGTCTCAGAAGGTTTTCCAGGGATGTGGACCCCCAAGCCCTCCAGCTGGACGAATTTCTCGT
TCCATCTCTGAAAGTGCCTTCAGTGCTCGCTTCAGACCACATCACCCGAGGAACGCCCAACCACAGCAGCTGGCACT
AGTTTGACCGACTGGTTACTGATGTCAAGGAGAACTGAAACAGGCCAAGAAATCTGGTCTCCCTTCCGAGCAAC
GTTTGCAACGATGAGAGGATGGCTGCAGGAACGGCAATGAGGATGACTGTTGGAATGGGAAAGGCAAAAGCAGGTAC
CTGTTTGCAGTGACAGGAATGGATTAGCCAACAGGGCAACAACCCAGAGGTCCAGGTTGACACCAGCAAAACCAGAC
ATACTGATCCTTCGTCAAATCATGGCTCTTCGAGTGATGACCAGCAAGATGAAGAATGCATACAATGGGAACGACGTG
GACTTCTTTGATATCAGTGATGAAAGTAGTGGAGAAGGAAGTGGAAAGTGGCTGTGAGTATCAGCAGTGCCCTTCAGAG
TTTGACTACAATGCCACTGACCATGCTGGGAAGAGTGCCAATGAGAAAGCCGACAGTGCTGGTGTCCGTCTCGGGCA
CAGGCCCTACCTCCTCACTGTCTTCTGCATCTTGTCTGCTTATGCAGAGAGAGTGGAGATTAATTCTCAAACCTCTGAG
AAAAAGTGTTCATCAAAAGTTAAAAGGCACCAAGTTATCACTTTTCTACCATCCTAGTGACTTTGCTTTTAAATGAA
TGGACAACAATGTACAGTTTTTACTATGTGGCACTGGTTTAAGAAGTGTGACTTTGTTTTCTCATTAGTTTTGGG
AGGAAAAGGGACTGTGCATTGAGTTGTTCTGCTCCCCAAACCATGTTAAACGTGGCTAACAGTGTAGGTACAGAA
CTATAGTTAGTTGTGCATTTGTGATTTTATCACTCTATTATTGTTTGTATGTTTTTCTCATTTCGTTTGTGGGTT
TTTTTTTCAAACGTGATCTCGCCTTGTTTCTTACAAGCAAACAGGGTCCCTTCTTGGCAGTAACATGTACGTATT
TCTGAAATATTAAATAGCTGTACAGAAGCAGGTTTTATTATCATGTTATCTATTAAAGAAAAAGCCCAAAAGC

FIGURE 24

MARFGLPALLCTLAVLSAALLAELKSKSCSEVRRLYVSKGFNKNDAPLHEINGDHLKICPQGST
CCSQEMEKEYSLQSKDDFKSVVSEQCNHLQAVFASRYKKFDEFFKELLENAEKSLNDMFVKTYGH
LYMQNSELFKDLFVELKRYVVGNNLEMLNDFWARLLERMFLVNSQYHFTDEYLECVSKYTE
QLKPFQDVPRKLKLVTRAFVAARTFAQGLAVAGDVVSKVSVVNPTAQCTHALLKMIYCSHCRGL
VTVKPCYNYCSNIMRGCLANQGDLD FEWNNFIDAMLMVAERLEGPFNIESVMDPIDVKISDAIMN
MQDNSVQVSQKVFQCGPPKPLPAGRISRISSESASFARFRPHHPEERPTTAAGTSLDRLVTDVK
EKLKQAKKFWSSLPSNVCNDERMAAGNGNEDDCWNGKGKSRYLFAVTGNGLANQGNPEVQVDT
KPDILILRQIMALRVMTSKMKNAYNGNDVDFDISDESSGEGSGSGCEYQQCPSEFDYNATDHAG
KSANEKADSAGVRPGAQAYLLTVFCILFLVMQREWR

Important features:

Signal peptide:

amino acids 1-22

ATP/GTP-binding site motif A (P-loop).

amino acids 515-524

N-glycosylation site.

amino acids 514-518

Glycosaminoglycan attachment sites.

amino acids 494-498, 498-502

N-myristoylation sites.

amino acids 63-69, 224-230, 276-282, 438-444, 497-503, 531-537

Glypicans proteins.

amino acids 54-75, 105-157, 238-280, 309-346, 423-460, 468-506

FIGURE 25

CTCGCCCTCAAATGGGAACGCTGGCCTGGGACTAAAGCATAGACCACCAGGCTGAGTATCCTGAC
CTGAGTCATCCCCAGGGATCAGGAGCCTCCAGCAGGGAACCTTCCATTATATTCTTCAAGCAACT
TACAGCTGCACCGACAGTTGCCGATGAAAGTTCTAATCTCTTCCCTCCTCCTGTTGCTGCCACTAA
TGCTGATGTCCATGGTCTCTAGCAGCCTGAATCCAGGGGTGCCCAGAGGCCACAGGGACCGAGGC
CAGGCTTCTAGGAGATGGCTCCAGGAAGGCGCCAAGAATGTGAGTGCAAAGATTGGTTCCTGAG
AGCCCCGAGAAGAAAATTTCATGACAGTGTCTGGGCTGCCAAAGAAGCAGTGCCCCTGTGATCATT
TCAAGGGCAATGTGAAGAAAACAAGACACCAAGGCACCACAGAAAGCCAAACAAGCATTCCAGA
GCCTGCCAGCAATTTCTCAAACAATGTCAGCTAAGAAGCTTTGCTCTGCCTTTGTAGGAGCTCTG
AGCGCCCACTCTTCCAATTAAACATTCTCAGCCAAGAAGACAGTGAGCACACCTACCAGACACTC
TTCTTCTCCCACCTCACTCTCCCCTGTACCCACCCCTAAATCATTCCAGTGCTCTCAAAAAGCA
TGTTTTTCAAGATCATTTTGGTTTGTGCTCTCTCTAGTGTCTTCTTCTCTCGTCAGTCTTAGCCT
GTGCCCTCCCCTTACCCAGGCTTAGGCTTAATTACCTGAAAGATTCCAGGAACTGTAGCTTCCT
AGCTAGTGTCAATTTAACCTTAAATGCAATCAGGAAAGTAGCAAACAGAAGTCAATAAATATTTTT
AAATGTCAAAAAAAAAAAAAAAAAAA

FIGURE 26

MKVLISLSSLLLLPLMLMSMVSSSLNPGVARGHRDRGQASRRWLQEGGQECECKDWFLRAPRRKFM
TVSGLPKKQCPCDHFKGNVKKTRHQRHHRKPNKHSRACQQFLKQCQLRSFALPL

Important features:

Signal peptide:

amino acids 1-22

N-myristoylation sites.

amino acids 27-33, 46-52

10063551

FIGURE 27

GGACGCCAGCGCCTGCAGAGGCTGAGCAGGGAAAAAGCCAGTGCCCCAGCGGAAGCACAGCTCAG
AGCTGGTCTGCCATGGACATCCTGGTCCCCTCCTGCAGCTGCTGGTGTGCTTCTTACCCTGCC
CCTGCACCTCATGGCTCTGCTGGGCTGCTGGCAGCCCCGTGTCAAAAGCTACTTCCCCTACCTGA
TGGCCGTGCTGACTCCCAAGAGCAACCGCAAGATGGAGAGCAAGAAACGGGAGCTCTTCAGCCAG
ATAAAGGGGCTTACAGGAGCCTCCGGGAAAGTGGCCCTACTGGAGCTGGGCTGCGGAACCGGAGC
CACTTTTCAGTTCTACCCACCGGGCTGCAGGGTCACCTGCCTAGACCCAAATCCCCACTTTGAGA
AGTTCCTGACAAAGAGCATGGCTGAGAACAGGCACCTCCAATATGAGCGGTTTGTGGTGGCTCCT
GGAGAGGACATGAGACAGCTGGCTGATGGCTCCATGGATGTGGTGGTCTGCACTCTGGTGTGTG
CTCTGTGCAGAGCCCAAGGAAGGTCTGCAGGAGGTCCGGAGAGTACTGAGACCGGGAGGTGTGC
TCTTTTTCTGGGAGCATGTGGCAGAACCATATGGAAGCTGGGCCTTCATGTGGCAGCAAGTTTTTC
GAGCCACCTGGAACACATTGGGGATGGCTGCTGCCTCACCAGAGAGACCTGGAAGGATCTTGA
GAACGCCCAGTTCTCCGAAATCCAAATGGAACGACAGCCCCCTCCCTTGAAGTGGCTACCTGTTG
GGCCCCACATCATGGGAAAGGCTGTCAAACAATCTTTCCCAAGCTCCAAGGCACTCATTTGCTCC
TTCCCCAGCCTCCAATTAGAACAAAGCCACCCACCAGCCTATCTATCTTCCACTGAGAGGGACCTTA
GCAGAATGAGAGAAGACATTGATGTACCACCTACTAGTCCCTCTCTCCCCAACCTCTGCCAGGGC
AATCTCTAACTTCAATCCCGCCTTCGACAGTGAAAAAGCTCTACTTCTACGCTGACCCAGGGAGG
AAACACTAGGACCCTGTTGTATCCTCAACTGCAAGTTTCTGGACTAGTCTCCCAACGTTTGCCTC
CCAATGTTGTCCCTTTCCCTTCGTTCCCATGGTAAAGCTCCTCTCGCTTTCTCCTGAGGCTACAC
CCATGCGTCTCTAGGAACTGGTCACAAAAGTCATGGTGCCTGCATCCCTGCCAAGCCCCCTGAC
CCTCTCTCCCCACTACCACCTTCTTCTGAGCTGGGGGCACCAGGGAGAATCAGAGATGCTGGGG
ATGCCAGAGCAAGACTCAAAGAGGCAGAGGTTTTGTTCTCAAATATTTTTTAATAAATAGACGAA
ACCACG

FIGURE 28

MDILVPLLQLLVLLLTPLPLHLMALLGCWQPLCKSYFPYLMAVLTPKSNRKMESKKRELFSSQIKGL
TGASGKVALLELGCCTGANFQFYPPGCRVTCCLDPNPHFEKFLTKSMAENRHLQYERFVVAPGEDM
RQLADGSMDVVVCTLVLCVQSPRKVLQEVRRVLRPGGVLFWEHVAEPYGSWAFMWQQVFPEPTW
KHIGDGCCLTRETWKDLNAQFSEIQMERQPPPLKWLPGPHIMGKAVKQSFPSKALICSFPSL
QLEQATHQPIYLPLRGT

Important features:

Signal peptide:

amino acids 1-23

Leucine zipper pattern.

amino acids 10-32

N-myristoylation sites.

amino acids 64-70, 78-84, 80-86, 91-97, 201-207

FIGURE 29

CAATGTTTGCCTATCCACCTCCCCAAGCCCCTTTACCTATGCTGCTGCTAACGCTGCTGCTGCT
GCTGCTGCTGCTTAAAGGCTCATGCTTGGAGTGGGGACTGGTCGGTGCCCAGAAAGTCTCTTCTG
CCACTGACGCCCCCATCAGGGATTGGGCCTTCTTTCCCCCTTCCTTTCTGTGCTCCTGCCTCAT
CGGCCTGCCATGACCTGCAGCCAAGCCCAGCCCCGTGGGGAAGGGGAGAAAGTGGGGGATGGCTTA
AGAAAGCTGGGAGATAGGGAACAGAAGAGGGTAGTGGGTGGGCTAGGGGGGCTGCCTTATTTAA
GTGGTTGTTTATGATTCTTATACTAATTTATACAAAGATATTAAGGCCCTGTTTATTAAAGAAATT
GTTCCCTTCCCCGTGTGTTCAATGTTTGTAAAGATTGTTCTGTGTAAATATGTCTTTATAATAAAC
AGTTAAAAGCTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 30

MLLLTLTLLLLLLKGSCLWGLVGAQKVSSATDAPIRDWAFFPPSFLCLLPHRPAMTCSQAQPRG
EGEKVGDG

Important features:

Signal peptide:

amino acids 1-15

Growth factor and cytokines receptors family:

amino acids 3-18

FIGURE 31

GTTTGAATTCCTTCAACTATACCCACAGTCCAAAAGCAGACTCACTGTGTCCCAGGCTACCAAGTT
CCTCCAAGCAAGTCATTTCCCTTATTTAACCGATGTGTCCCTCAAACACCTGAGTGCTACTCCCT
ATTTGCATCTGTTTTGATAAATGATGTTGACACCCCTCCACCGAATTCCTAAGTGGAATCATGTCGG
GAAGAGATACAATCCTTGGCCTGTGTATCCTCGCATTAGCCTTGTCTTTGGCCATGATGTTTACC
TTCAGATTCATCACCACCCTTCTGGTTCACATTTTCATTTTCATTGGTTATTTTGGGATTGTTGTT
TGTCTGCGGTGTTTTATGGTGGCTGTATTATGACTATACCAACGACCTCAGCATAGAATTGGACA
CAGAAAGGGAAAATATGAAGTGCCTGCTGGGGTTTGTCTATCGTATCCACAGGCATCACGGCAGTG
CTGCTCGTCTTGATTTTTGTTCTCAGAAAGAGAATAAAATTGACAGTTGAGCTTTTCCAAATCAC
AAATAAAGCCATCAGCAGTGCTCCCTTCTGCTGTTCAGCCACTGTGGACATTTGCCATCCTCA
TTTTCTTCTGGGTCTCTGGGTGGCTGTGCTGCTGAGCCCTGGGAACTGCAGGAGCTGCCAGGTT
ATGGAAGGCGGCCAAGTGGAATATAAGCCCCCTTTCGGGCATTCCGTACATGTGGTCGTACCATTT
AATTGGCCTCATCTGGACTAGTGAATTCATCCTTGCCTGCCAGCAAATGACTATAGCTGGGGCAG
TGGTTACTTGTATTTTCAACAGAAGTAAAAATGATCCTCCTGATCATCCCATCCTTTTCGTCTCTC
TCCATTCTCTTCTTCTACCATCAAGGAACCGTGTGAAAGGGTCATTTTTAATCTCTGTGGTGAG
GATTCCGAGAATCATTGTCTGTACATGCAAAACGCACTGAAAGAACAGCAGCATGGTGCATTGT
CCAGGTACCTGTTCCGATGCTGCTACTGCTGTTTCTGGTGCTTGACAAATACCTGCTCCATCTC
AACCAGAATGCATATACTACAACGCTATTAAATGGGACAGATTTCTGTACATCAGCAAAAGATGC
ATTCAAAATCTTGTTCCAAGAACTCAAGTCACCTTTACATCTATTAAGTCTTTGGAGACTTCATAA
TTTTTCTAGGAAAGGTGTTAGTGGTGTGTTTCACTGTTTTTGGAGGACTCATGGCTTTTAACTAC
AATCGGGCATTCCAGGTGTGGGCAGTCCCTCTGTTATTGGTAGCTTTTTTGCCTACTTAGTAGC
CCATAGTTTTTTATCTGTGTTTGAAACTGTGCTGGATGCACCTTTTCTGTGTTTTGCTGTGATC
TGGAACAAATGATGGATCGTCAGAAAAGCCCTACTTTATGGATCAAGAATTTCTGAGTTTCGTA
AAAAGGAGCAACAAATTAAACAATGCAAGGGCACAGCAGGACAAGCACTCATTAAAGGAATGAGGA
GGGAACAGAACTCCAGGCCATTGTGAGATAGATACCCATTTAGGTATCTGTACCTGGAAAACATT
TCCTTCTAAGAGCCATTTACAGAATAGAAGATGAGACCACTAGAGAAAAGTTAGTGAATTTTTTT
TTAAAAGACCTAATAAACCTATTCTTCCTCAAAA

FIGURE 32

MSGRTILGLCILALALSLAMMFTFRFITTLVHIFISLVILGLLFVCGVLWWLYDYDNDLSIE
LDTERENMKCVLGFAIVSTGITAVLLVLIFVLRKRIKLTVELFQITNKAISSAPFLLFQPLWTF
ILIFFWVLWVAVLLSLGTAGAAQVMEGGQVEYKPLSGIRYMWSYHLIGLIWTSEFILACQOMTIA
GAVVTCYFNRSKNPPDHPILSSLSILFFYHQGTVVKGSFLISVVRIPRIIVMYMQLKEQQHG
ALSRYLFRCCYCCFWCLDKYLLHLNQAYTTTAINGTDFCTSAKDAFKILSKNSSHFTSINCFGD
FIIIFLGKVLVVCFTVFGGLMAFNYNRAQVWAVPLLLVAFFAYLVAHSFLSVFETVLDALFLCFA
VDLETNDGSSEKPYFMDQEFLSFVKRSNKLNNARAQQDKHSRNEEGTELQAIVR

Important features:

Signal peptide:

amino acids 1-20

Putative transmembrane domains:

amino acids 35-54, 75-97, 126-146, 185-204, 333-350, 352-371

N-glycosylation sites.

amino acids 204-208, 295-299, 313-317

N-myristoylation sites.

amino acids 147-153, 178-184, 196-202, 296-275, 342-348

FIGURE 33

GTTCGATTAGCTCCTCTGAGAAGAAGAGAAAAGGTTCTTGACCTCTCCCTGTTTCTTCCTTAGA
 ATAATTTGTATGGGATTTGTGATGCAGGAAAGCCTAAGGGAAAAAGAATATTCATTCTGTGTGGT
 GAAAATTTTTTGAAGGAAATTCCTTCTTCAAACAAGGGTGTCTATTCTGATATTTATCAGGAC
 TGTTGTTCTCTACTATGAAGGCATCTGTTATTGAAATGTTCTTGTGTTTGTCTGCTGACTGGAGTAC
 ATTCAAACAAAGAAACGGCAAAGAAGATTAAAGGCCCAAGTTCCTGTGCCTCAGATCAACTGC
 GATGTCAAAGCCGGAAAGATCATCGATCCTGAGTTCATTGTGAAATGTCCAGCAGGATGCCAAGA
 CCCCAAATACCATGTTTATGGCACTGACGTGTATGCATCCTACTCCAGTGTGTGTGGCGCTGCCG
 TACACAGTGGTGTGCTTGATAATTGAGGAGGAAAAATACTTGTTCGGAAGGTTGCTGGACAGTCT
 GGTTCACAAAGGGAGTTATTCCAACGGTGTCCAATCGTTATCCCTACCACGATGGAGAGAATCCTT
 TATCGTCTTAGAAAGTAAACCCAAAAAGGGTGTAACTACCCATCAGCTCTTACATACTCATCAT
 CGAAAGTCCAGCTGCCCAAGCAGGTGAGACCACAAAAGCCTATCAGAGGCCACCTATCCAGGG
 ACACTGCACAGCCGGTCACTCTGATGCAGCTTCTGGCTGTCACTGTAGCTGTGGCCACCCCCAC
 CACCTTGCCAAAGGCCATCCCCTTCTGCTGCTTCTACCACCAGCATCCCCAGACCACAATCAGTGG
 GCCACAGGAGCCAGGAGATGGATCTCTGGTCCACTGCCACCTACACAAGCAGCCAAAACAGGCCG
 AGAGCTGATCCAGGTATCCAAAGGCAAGATCCTTCAGGAGCTGCCTTCCAGAAACCTGTTGGAGC
 GGATGTCAGCCTGGGACTTGTTCAAAAGAAGAATTGAGCACACAGTCTTGGAGCCAGTATCCC
 TGGGAGATCCAACTGCAAAATTGACTTGTCTGTTTTTAATTGATGGGAGCACCAGCATTGGCAAA
 CGGCGATTCCGAATCCAGAAGCAGCTCCTGGCTGATGTTGCCAAGCTCTTGACATTGGCCCTGC
 CGGTCCACTGATGGGTGTTGTCCAGTATGGAGACAACCTGCTACTCACTTTAACCTCAAGACAC
 ACACGAATTTCTGAGATCTGAAGACAGCCATAGAGAAAATTACTCAGAGAGGAGACTTTCTAAT
 GTAGGTCGGGCCATCTCCTTTTGTGACCAAGAAGTCTTTTCCAAAGCCAATGGAAACAGAAGCGG
 GGCTCCCAATGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT
 GACTTGGGAGAGAGTCAGGAATCAACATTTTCTTCATCACCATTGAAGGTGCTGCTGAAAATGAG
 AAGCAGTATGTGGTGGAGGCCAATTTGCAACAAGGCCGTGTGCAGAACAAACGGCTTCTACTC
 GCTCCACGTGCAGAGCTGGTTTGGCCTCCACAAGACCCTGCAGCCTCTGGTGAAGCGGGTCTGCG
 AACTGACCGCCTGGCCTGCAGCAAGACCTGCTTGAACCTGGCTGACATTGGCTTCGTCATCGAC
 GGCTCCAGCAGTGTGGGGACGGGCAACTTCCGCACCGTCTCCAGTTTGTGACCAACCTCACCAA
 AGAGTTTGAGATTCCGACACGGACACGCGCATCGGGCCGTGCAGTACACCTACGAACAGCGGC
 TGGAGTTTGGGTTTCGACAAGTACAGCAGCAAGCCTGACATCCTCAACGCCATCAAGAGGGTGGGC
 TACTGGAGTGGTGGCACCAGCACGGGGGCTGCCATCAACTTCGCCCTGGAGCAGCTCTTCAAGAA
 GTCCAAGCCCAACAAGAGGAAGTTAATGATCCTCATCACCAGCGGAGGTCTACGACGACGTCC
 GGATCCCAGCCATGGCTGCCATCTGAAGGGAGTGATCACCTATGCGATAGGCGTTGCCTGGGCT
 GCCAAGAGGAGCTAGAAGTCATTGCCACTCACCCCGCCAGAGACCACTCCTTCTTTGTGGACGA
 GTTTGACAACCTCCATCAGTATGTCCCCAGGATCATCCAGAACATTTGTACAGAGTTCAACTCAC
 AGCCTCGGAAC**TGA**ATTGAGAGCAGGCAGAGCACCAGCAAGTGCTGCTTTACTAACTGACGTGTT
 GGACCACCCACCGCTTAATGGGGCACGCACGGTGCATCAAGTCTTGGGCAGGGCATGGAGAAAC
 AAATGTCTTGTTATTATTCTTTGCCATCATGCTTTTTCATATTCCAAAACCTTGAGATTACAAAGA
 TGATCACAACGTATAGAATGAGCCAAAAGGCTACATCATGTTGAGGGTGTGAGATTTTACAT
 TTTGACAATTGTTTTCAAAATAAATGTTCCGAATACAGTGCAGCCCTTACGACAGGCTTACGTAG
 AGCTTTTGTGAGATTTTAAAGTTGTTATTTCTGATTTGAACTCTGTAACCTCAGCAAGTTTCAT
 TTTTGTGATGACAATGTAGGAATTGCTGAATTAATGTTTGAAGGATGAAAAATAAAAAA
 AA
 AA

FIGURE 34

MRTVVLTMKASVIEMFLVLLVTGVHSNKETAKKIKRPKFTVPQINCDVKAGKIIDPEFIVKCPAG
CQDPKYHVYGTDVYASYSSVCGAAVHSGVLDNSGGKILVRKVAGQSGYKGSYNGVQSLSLPRWR
ESFIVLESKPKKGVTPSALTYSSSKSPAAQAGETTKAYQRPPPIPGTTAQPVTLMQLLAVTVAVA
TPTTLRPPSPAASTTSIPRPQSVGHRSEQEMDLWSTATYTSSQNRPRADPGIQRQDPSGAAFQKP
VGADVSLGLVPKEELSTQSLEPVSLGDPNCKIDLSFLIDGSTSIGKRRFRIQKQLLADVAQALDI
GPAGPLMGVVQYGDNPATHFNLKTHTNRDLKTAIEKITQRGGLSNVGRAISFVTKNFFSKANGN
RSGAPNVVVVMVDGWPTDKVEEASRLARESGINIFFITIEGAAENEKQYVVEPNFANKAVCRTNG
FYSLHVQSWFGLHKTLPVLRVCDTDRACSKTCLNSADIGFVIDGSSSVGTGNFRTVLQFVTN
LTKEFEISDTRIGAVQYTYEQRLEFGFDKYSSKPDILNAIKRVGYWSGGTSTGAAINFALEQL
FKKSKPNKRKLMILITDGRSYDDVRIPAMAAHLKGVITYAIGVAAWAAQEELEVIATHPARDHSFF
VDEFNLHQYVPRIIQNICTEFNSQPRN

Important features:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 181-200

N-glycosylation sites.

amino acids 390-394, 520-524

N-myristoylation sites.

amino acids 23-29, 93-99, 115-121, 262-268, 367-373, 389-395,
431-437, 466-472, 509-515, 570-576, 571-577, 575-581, 627-633

Amidation site.

amino acids 304-308

[illegible]

	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100	2101	2102	2103	2104	2105	2106	2107	2108	2109	2110	2111	2112	2113	2114	2115	2116	2117	2118	2119	2120	2121	2122	2123	2124	2125	2126	2127	2128	2129	2130	2131	2132	2133	2134	2135	2136	2137	2138	2139	2140	2141	2142	2143	2144	2145	2146	2147	2148	2149	2150	2151	2152	2153	2154	2155	2156	2157	2158	2159	2160	2161	2162	2163	2164	2165	2166	2167	2168	2169	2170	2171	2172	2173	2174	2175	2176	2177	2178	2179	2180	2181	2182	2183	2184	2185	2186	2187	2188	2189	2190	2191	2192	2193	2194	2195	2196	2197	2198	2199	2200	2201	2202	2203	2204	2205	2206	2207	2208	2209	2210	2211	2212	2213	2214	2215	2216	2217	2218	2219	2220	2221	2222	2223	2224	2225	2226	2227	2228	2229	2230	2231	2232	2233	2234	2235	2236	2237	2238	2239	2240	2241	2242	2243	2244	2245	2246	2247	2248	2249	2250	2251	2252	2253	2254	2255	2256	2257	2258	2259	2260	2261	2262	2263	2264	2265	2266	2267	2268	2269	2270	2271	2272	2273	2274	2275	2276	2277	2278	2279	2280	2281	2282	2283	2284	2285	2286	2287	2288	2289	2290	2291	2292	2293	2294	2295	2296	2297	2298	2299	2300	2301	2302	2303	2304	2305	2306	2307	2308	2309	2310	2311	2312	2313	2314	2315	2316	2317	2318	2319	2320	2321	2322	2323	2324	2325	2326	2327	2328	2329	2330	2331	2332	2333	2334	2335	2336	2337	2338	2339	2340	2341	2342	2343	2344	2345	2346	2347	2348	2349	2350	2351	2352	2353	2354	2355	2356	2357	2358	2359	2360	2361	2362	2363	2364	2365	2366	2367	2368	2369	2370	2371	2372	2373	2374	2375	2376	2377	2378	2379	2380	2381	2382	2383	2384	2385	2386	2387	2388	2389	2390	2391	2392	2393	2394	2395	2396	2397	2398	2399	2400	2401	2402	2403	2404	2405	2406	2407	2408	2409	2410	2411	2412	2413	2414	2415	2416	2417	2418	2419	2420	2421	2422	2
--	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	---

FIGURE 36

MASALWTVLPSRMSLRSLKWSLLLLSLLSFFVMWYLSLPHYNVIERVNWMYFYEYEPYRQDFHF
TLREHSNCSHQNPFLVILVTSHPSDVKARQAIRVTWGEKKSWWGYEVLTFLLGQEAEKEDKMLA
LSLEDEHLLYGDIIRQDFLDTYNNLTTLKTIMAFRWVTEFCPNAKYVMKTDTDVFINTGNLVKYL
NLNHSEKFFTGYPLIDNYSYRGFYQKTHISYQEYPFKVFPYCSGLGYIMSRDLVPRIYEMMGHV
KPIKFEDVYVGICLNLKVNHIHIPEDTNLFFLYRIHLDVCQLRRVIAAHGFSSKEIITFWQVMLR
NTTCHY

Important features:

Type II transmembrane domain:

amino acids 20-39

N-glycosylation sites.

amino acids 72-76, 154-158, 198-202, 212-216, 326-330

Glycosaminoglycan attachment site.

amino acids 239-243

Ly-6 / u-PAR domain proteins.

amino acids 23-37

N-myristoylation site.

amino acids 271-277

FIGURE 37

CGCTCGGGCACCAGCCGCGGCAAGGATGAGAGCTGGGTTGCTGGACGAGTTGGGGCTCAGTTTTCTTCAGCTCCTTCTCATC
TCGTCTCTGCCAAGAGAGTACACAGTCATTAATGAAGCCTGCCCTGGAGCAGAGTGGAAATATCATGTGTCGGGAGTGCTGTG
AATATGATCAGATTGAGTGCCTCTGCCCGGAAAGAGGGAAGTCGTGGGTTATACCATCCCTTGCTGCAGGAATGAGGAGAA
TGAGTGTGACTCCTGCCGTGATCCACCCAGGTTGTACCATCTTTGAAAAGTGCAGAGCTGCCGAAATGGCTCATGGGGGGT
ACCTTGAGTACTTCTATGTGAAGGGGTTCTACTGTGCAGAGTGCCGAGCAGGCTGGTACGGAGGAGACTGCATGCGATGTG
GCCAGGTTCTGCGAGCCCCAAGGGTGCAGATTTTGTGAAAAGCTATCCCTAAATGCTCACTGTGAATGGACCATTTCATGC
TAAACCTGGGTTTGTCAATCAACTAAGATTTGTCTGTGAGTCTGGAGTTTGGACTACATGTGCCAGTATGACTATGTTGAG
GTTCTGTGATGGAGACAACCGCATGCCAGATCATCAAGCGTGTCTGTGGCAACGAGCGGCCAGCTCCTATCCAGAGCATAG
GATCCTCACTCCACGTCCTCTTCCACTCCGATGGCTCCAAGAATTTGACGGTTTCCATGCCATTATGAGGAGATCACAGC
ATGCTCCTCATCCCTTGTTTCCATGACGGCAGCGTCCCTTGACAAGGCTGGATCTTACAAGTGTGCTGCTTGGCAGGC
TATACTGGGCAGCGCTGTGAAAATCTCCTTGAGAAAAGAACTGCTCAGACCTGGGGGCCAGTCAATGGGTACCAGAAAA
TAACAGGGGGCCCTGGGCTTATCAACGGACGCCATGCTAAAATGGCACCGTGGTGTCTTTCTTTTGTAACTCCTATGT
TCTTAGTGGCAATGAGAAAAGAACTTGCCAGCAGAAATGGAGAGTGGTCAGGAAACAGCCCATCTGCATAAAAGCCTGCCGA
GAACCAAGATTTAGACCTGGTGAGAGGAGAGTTCTTCCGATGCAGGTTCACTCAAGGGAGACACCATTACACAGCTAT
ACTCAGCGGCTTACGCAAGCAGAACTGCAGAGTGGCCTACCAAGAGCCAGCCCTTCCCTTTGGAGATCTGCCCATGGG
ATACCAACATCTGCATACCCAGCTCCAGTATGAGTGCATCTACCCCTTCTACCGCCGCTGGGCAGCAGCAGGAGGACATGT
CTGAGGACTGGGAAGTGGAGTGGCGGGGACCATCTGCATCCCTATCTGCGGGAAAATGAGAACATCACTGCTCCAAAGA
CCCAAGGGTTGCGCTGGCGTGGCAGGCAGCCATCTACAGGAGGACCAGCGGGTGCATGACGGCAGCCTACACAAGGGAGC
GTGGTTCTTAGTCTGCAGCGGTGCCCTGGTGAATGAGCGCACTGTGGTGGTGGCTGCCACTGTGTACTGACCTGGGGAAG
GTCACCATGATCAAGACAGCAGACCTGAAGTTGTTTTGGGAAATTTACCGGGATGATGACCGGGATGAGAAGACCATCC
AGAGCCTACAGATTTCTGCTATCATCTGCATCCCAACTATGACCCCATCTGCTTGATGCTGACATCGCCATCCTGAAGCT
CCTAGACAAGGCCGATCAGACCCGAGTCCAGCCATCTGCCCTGCTGCCAGTGGGATCTCAGCATTCCTTCCAGGAG
TCCCACATCACTGTGGCTGGCTGGAATGTCTTGGCAGAGTGGAGGCCCTGGCTTCAAGAACGACACACTGCGCTCTGGGG
TGCTCAGTGTGGTGGACTCGCTGCTGTGTGAGGAGCAGCATGAGGACCATGGCATCCAGTGAGTGTCACTGATAACATGTT
CTGTGCCAGCTGGGAACCCACTGCCCTTCTGATATCTGCATCTGCAGAGACAGGAGGCATCGCGGCTGTGTCTTCCCGGA
CGAGCATCTCCTGAGCCACGCTGGCATCTGATGGGACTGCTCAGCTGGAGCTATGATAAAACATGCAGCCACAGGCTCTCCA
CTGCCCTTACCAAGGTGCTGCCCTTTAAAGACTGGATTGAAAGAAATATGAAATGAACCATGCTCATGCACTCCTTGAGAAG
TGTTTCTGATATCCGCTCTGTACGTGTGTCATTGCGTGAAGCAGTGTGGGCTGAAGTGTGATTGGCCTGTGAACCTGGCT
GTGCCAGGGCTTCTGACTTCAGGGACAAAACCTCAGTGAAGGGTGAAGTAGCTCCATTGCTGGTAGGCTGATGCCGCTCCA
CTACTAGGACAGCCAATTGGAAGATGCCAGGGCTTGAAGAAGTAAGTTTCTTCAAGAGACCATATACAAAACCTCTCCA
CTCCACTGACCTGGTGGCTTCTCCCAACTTTTCAAGTTATACGAATGCCATCAGCTTGACCGGGAAGATCTGGGCTTCATGAG
GCCCCTTTGGAGGCTCTCAAGTTCTAGAGAGCTGCCGTGGGACAGCCAGGGCAGCAGAGCTGGGATGTGGTGCATGCCCTT
TGTGTACATGGCCACAGTACAGTCTGGTCTTTTCTTCCCCATCTCTTGTACACATTTTAAATAAATAAGGGTTGGCTTCT
GAACTACAAAAA
AA

FIGURE 38

MELGCWTQLGLTFLQLLLISSLPREYTVINEACPGAENIMCRECCEYDQIECVCPGKREVVGYT
IPCCRNEENECDSCLIHPGCTIFENCKSCRNGSWGGLDDFYVKGIFYCAECRAGWYGGDCMRGQ
VLRAPKGQILLESYPLNAHCEWTIHAKPGFVIQLRFVMLSLEFDYMCQYDYVEVRDGDNRDQII
KRVCGNERPAPIQSIGSSSLHVLHSDGSKNFDGFHAIYEEITACSSSPCFHDGTCVLDKAGSYKC
ACLAGYTGQRCENLLEERNCSDFGGPVNGYQKITGGPGLINGRHAKIGTVVSFFCNSYVLSGNE
KRTCQQNGEWSGKQPICIKACREPKISDLVRRRVLPMPVQVSRETPLHQLYSAAFSKQKLQSAPTK
KPALPFGDLPMGYQHLHTQLQYECISPFYRRLGSSRRRTCLRTGKWSGRAPSCIPICGKIENITAP
KTQGLRWPDQAAIYRRTSGVHDGSLHKGAWFLVCSGALVNERTVVVAHCVTDLGKVTMIKTADL
KVVLGKFYRDDDRDEKTIQSLQISAILHPNYDPILLDADIAILKLLDKARISTRVQPICLAASR
DLSTSFQESHITVAGWNVLADVRSPGFKNDTLRSGVSVVDSLLCEEQHEDHGIPVSVTDNMFCA
SWEPTAPSDICTAETGGIAAVSFPGRASPEPRWHLMLVSWSYDKTCSHRLSTAFTKVLFPKDWI
ERNMK

Important features of the protein:

Signal peptide:

amino acids 1-23

EGF-like domain cysteine pattern signature.

amino acids 260-272

N-glycosylation sites.

amino acids 96-100, 279-283, 316-320, 451-455, 614-618

N-myristoylation sites.

amino acids 35-41, 97-103, 256-262, 284-290, 298-304, 308-314,
474-480, 491-497, 638-644, 666-672

Amidation site.

amino acids 56-60

Serine proteases, trypsin family.

amino acids 489-506

CUB domain proteins profile.

amino acids 150-167

[illegible][illegible]

FIGURE 40

MKALLLLVLPWLSPANYIDNVGNLHFLYSELCKGASHYGLTKDRKRRSQDGCPCDGCASLTATAPS
PEVSAAATISLMTDEPGLDNPAYVSSAEDGQPAISPVDSGRSNRTRARPFERSTIRSRSFKKINR
ALSVLRRTKSGSAVANHADQGRESENTTAPEVFPRLYHLIPDGEITSIKINRVDPSESLSIRLV
GGSETPLVHIIIQHIYRDGVIARDGRLLPGDIILKVNMDISNVPHNYAVRLLRQPCQVLWLTVM
REQKFRSRNNGQAPDAYRPRDDSFHVILNKSSPEEQLGIKLVRKVDEPGVFI FNVLDGGVAYRHG
QLEENDRVLAINGHDLRYGSPESAHLIQASERRVHLVVSQRQRSPDIFQEAGWNSNGSWSPG
PGRSNTPKPLHPTITCHEKVNIQKDPGESLGMTVAGGASHREWDLPYVISVEPGGVISRDGR
IKTGDILLNVGDVELTEVSRSEAVALLKRTSSSIVLKALEVKEYEPQEDCSSPAALDSNHNMAPP
SDWSPSWMMWLELPRCLYNCKDIVLRRTAGSLGFCIVGGYEEYNGNKPFFIKSIVEGTPAYNDG
RIRCGDILLAVNGRSTSGMIHACLARLLKELKGRITLTIVSWPGTFL

Important features:

Signal peptide:

amino acids 1-15

N-glycosylation sites.

amino acids 108-112, 157-161, 289-293, 384-388

Tyrosine kinase phosphorylation sites.

amino acids 433-441, 492-500

N-myristoylation sites.

amino acids 51-57, 141-147, 233-239, 344-350, 423-429, 447-453,
467-473, 603-609

FIGURE 41

ACCAGGCATTGTATCTTCAGTTGTCATCAAGTTCGCAATCAGATTGGAAAAGCTCAACTTGAAGCTTT
CTTGCCCTGCAGTGAAGCAGAGAGATAGATATTATTACGTAATAAAAAACATGGGCTTCAACCTGACT
TTCCACCTTTCTACAAATTCGATTACTGTTGCTGTTGACTTTGTGCCTGACAGTGGTTGGGTGGGC
CACCAGTAACTACTTCGTGGGTGCCATTCAAGAGATTCTTAAAGCAAAGGAGTTCATGGCTAATTTCC
ATAAGACCCTCATTTTGGGGAAGGGAATACTCTGACTAATGAAGCATCCACGAAGAAGGTAGAACTT
GACAACTGTCCTTCTGTGTCCTTACCTCAGAGGCCAGAGCAAGCTCATTTTCAAACCAGATCTCAC
TTTGAAGAGGTACAGGCAGAAAATCCCAAAGTGTCCAGAGGCCGGTATCGCCCTCAGGAATGTAAAG
CTTTACAGAGGGTCGCCATCCTCGTTCCCCACCGGAACAGAGAGAAACACCTGATGTACCTGCTGGAA
CATCTGCATCCCTTCCTGCAGAGGCAGCAGCTGGATTATGGCATCTACGTCATCCACCAGGCTGAAGG
TAAAAAGTTTAATCGAGCCAACTCTTGAATGTGGGCTATCTAGAAGCCCTCAAGGAAGAAAATTGGG
ACTGCTTTATATTCCACGATGTGGACCTGGTACCCGAGAATGACTTTAACCTTTACAAGTGTGAGGAG
CATCCCAAGCATCTGGTGGTTGGCAGGAACAGCACTGGGTACAGGTTACGTTACAGTGGATATTTTGG
GGGTGTTACTGCCCTAAGCAGAGAGCAGTTTTTCAAGGTGAATGGATTCTCTAACAACACTACTGGGGAT
GGGGAGGCGAAGACGATGACCTCAGACTCAGGGTTGAGCTCCAAAGAATGAAAATTTCCCGGCCCTG
CCTGAAGTGGGTAAATATACAATGGTCTTCCACACTAGAGACAAAGGCAATGAGGTGAACGCAGAACG
GATGAAGCTCTTACACCAAGTGTACGAGTCTGGAGAACAGATGGGTGAGTAGTTGTTCTTATAAAT
TAGTATCTGTGGAACACAATCCTTTATATATCAACATCACAGTGGATTTCTGGTTTGGTGCATGACCC
TGGATCTTTTGGTGATGTTTGAAGAACTGATTCTTTGTTTGCAATAATTTTGGCCTAGAGACTTCAA
ATAGTAGCACACATTAAGAACCTGTTACAGCTCATTGTTGAGCTGAATTTTCTCTTTTGTATTTTCT
TAGCAGAGCTCCTGGTGATGTAGACTATAAACAGTTGTAACAAGACAGCTTCTTAGTCATTTTGAT
CATGAGGGTTAAATATTGTAATATGGATACTTGAAGGACTTTATATAAAAGGATGACTCAAAGGATAA
AATGAACGCTATTTGAGGACTCTGCTTGAAGGAGATTTATTTAAATTGGAAGTAATATATTATGGGAT
AAAAGGCCACAGGAAATAAGACTGCTGAATGTCTGAGAGAACCAGAGTTGTTCTCGTCCAAGGTAGAA
AGGTACGAAGATACAATACTGTTATTATTATCCTGTACAATCATCTGTGAAGTGGTGGTGTGAGGT
GAGAAGGCGTCCACAAAAGAGGGGAGAAAAGGCGACGAATCAGGACACAGTGAACCTGGGAATGAAGA
GGTAGCAGGAGGGTGGAGTGTGCGCTGCAAAGGCAGCAGTAGCTGAGCTGGTTGCAGGTGCTGATAGC
CTTCAGGGGAGGACCTGCCAGGTATGCCCTCCAGTGATGCCACCAGAGAATACATTCTCTATTAGT
TTTTAAAGAGTTTTTGTAAAATGATTTTGTACAAGTAGGATATGAATTAGCAGTTTACAAGTTTACAT
ATTAATAATAATAATATGTCTATCAAATACCTCTGTAGTAAATGTGAAAAGCAAAA

FIGURE 42

MGFNLT FHLSYKFRLLLLLTLCLTVVGWATSNYFVGAIQEIPKAKEFMANFHKTLILGKGKTLTN
EASTKKVELDNCPSVSPYLRGQSKLIFKPDLTLEEVQAENPKVSRGRYRPQECKALQRVAILVPH
RNREKHLMYLLEHLHPFLQRQQLDYGIYVIHQAEKKFNRAKLLNVGYLEALKEENWDCFI FHDV
DLVPENDFNLYKCEEHPKHLVVGRNSTGYRLRYSGYFGGVTALSREQFFKVNGFSNNYWGWGGE
DDLRLRVELQRMKISRPLPEVGKYTMVFHTRDKGNEVNAERMKLLHQVSRVWRTDGLSSCSYKLV
SVEHNPLYINITVDFWFGA

Important features:

Signal peptide:

amino acids 1-27

N-glycosylation sites.

amino acids 4-8, 220-224, 335-339

Xylose isomerase proteins.

amino acids 191-202

FIGURE 43

GCTCAAGACCCAGCAGTGGGACAGCCAGACAGACGGCACGATGGCACTGAGCTCCCAGATCTGGG
CCGCTTGCCTCCTGCTCCTCCTCCTCGCCAGCCTGACCAGTGGCTCTGTTTTCCACAACAG
ACGGGACAAC TTGCAGAGCTGCAACCCCAGGACAGAGCTGGAGCCAGGGCCAGCTGGATGCCCAT
GTTCCAGAGGCGAAGGAGGCGAGACACCCACTTCCCCATCTGCATTTCTGCTGCGGCTGCTGTC
ATCGATCAAAGTGTGGGATGTGCTGCAAGACG TAGAACCTACCTGCCCTGCCCCGTCCCCTCCC
TTCTTATTTATTCCTGCTGCCCCAGAACATAGGTCTTGAATAAAATGGCTGGTTCTTTTGT
TCCAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 44

MALSSQIWAACLLLLLLASLTSGSVFPQQTGQLAELQPQDRAGARASWMPMFQRRRRRDTHFPI
CIFCCGCCHRSKCGMCCKT

Important features:

Signal peptide:

amino acids 1-24

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 58-59

N-myristoylation site.

amino acids 44-50

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 1-12

FIGURE 45

GTGGCTTCATTTTCAGTGGCTGACTTCCAGAGAGCAATATGGCTGGTTCCCCAACATGCCTCACCC
TCATCTATATCCTTTGGCAGCTCACAGGGTCAGCAGCCTCTGGACCCGTGAAAGAGCTGGTCGGT
TCCGTTGGTGGGGCCGTGACTTTCCCCCTGAAGTCCAAGTAAAGCAAGTTGACTCTATTGTCTG
GACCTTCAACACAACCCCTCTGTACCATACAGCCAGAAGGGGGCACTATCATAGTGACCCAAA
ATCGTAATAGGGAGAGAGTAGACTTCCCAGATGGAGGCTACTCCCTGAAGCTCAGCAAACCTGAAG
AAGAATGACTCAGGGATCTACTATGTGGGGATATACAGCTCATCACTCCAGCAGCCCTCCACCCA
GGAGTACGTGCTGCATGTCTACGAGCACCTGTCAAAGCCTAAAGTCACCATGGGTCTGCAGAGCA
ATAAGAATGGCACCTGTGTGACCAATCTGACATGCTGCATGGAACATGGGGAAGAGGATGTGATT
TATACCTGGAAGGCCCTGGGGCAAGCAGCCAATGAGTCCCATAATGGGTCCATCCTCCCCATCTC
CTGGAGATGGGGAGAAAGTGATATGACCTTCATCTGCGTTGCCAGGAACCCCTGTCAGCAGAACT
TCTCAAGCCCCATCCTTGCCAGGAAGCTCTGTGAAGGTGCTGCTGATGACCCAGATTCTCCATG
GTCTCCTGTGTCTCCTGTTGGTGCCCTCCTGCTCAGTCTCTTTGTACTGGGGCTATTTCTTTG
GTTTCTGAAGAGAGAGAGACAAGAAGAGTACATTGAAGAGAAGAAGAGAGTGGACATTTGTCGGG
AAACTCCTAACATATGCCCCCATTCTGGAGAGAACACAGAGTACGACACAATCCCTCACACTAAT
AGAACAATCCTAAAGGAAGATCCAGCAAATACGGTTTACTCCACTGTGGAAATACCGAAAAAGAT
GGAAAAATCCCCACTCACTGCTCAGGATGCCAGACACACCAAGGCTATTTGCCTATGAGAATGTTA
TCTAGACAGCAGTGCCTCCCTAAGTCTCTGCTCA

FIGURE 46

MAGSPTCLTLIYILWQLTGSAAAGPVKELVGSVGGAVTFPLKSKVKQVDSIVWTFNTTPLVTIQP
EGGTIIIVTQNRNRERVDFFPDGGYSLKLSKLKKNDSGIYYVGIYSSSLQQPSTQEYVLHVYEHLSK
PKVTMGLQSNKNGTCVTNLTCMEHGEEDVIYTWKALGQAANESHNGSILPISWRWGESDMTFIC
VARNPVSRNFSSPILARKLCEGAADDPDSSMVLLCLLLVPLLLSLFVLGLFLWFLKRERQEEYIE
EKKRVDICRETPNICPHSGENTYDTIPHTNRTILKEDPANTVYSTVEIPKKMENPHSLLTMPDT
PRLFAYENVI

Important features:

Signal peptide:

amino acids 1-22

Transmembrane domain:

amino acids 224-250

Leucine zipper pattern.

amino acids 229-251

N-glycosylation sites.

amino acids 98-102, 142-146, 148-152, 172-176, 176-180, 204-208,
291-295

FIGURE 47

GGCTCGAGCGTTTCTGAGCCAGGGGTGACCATGACCTGCTGCGAAGGATGGACATCCTGCAATGG
ATTCAGCCTGCTGGTTCTACTGCTGTTAGGAGTAGTTCTCAATGCGATACCTCTAATTGTCAGCT
TAGTTGAGGAAGACCAATTTCTCAAACCCCATCTCTTGCTTTGAGTGGTGGTCCCAGGAATT
ATAGGAGCAGGTCTGATGGCCATTCCAGCAACAACAATGTCCTTGACAGCAAGAAAAAGAGCGTG
CTGCAACAACAGAACTGGAATGTTTCTTTCATCATTTTTTCAGTGTGATCACAGTCATTGGTGCTC
TGTATTGCATGCTGATATCCATCCAGGCTCTCTTAAAAGGTCCTCTCATGTGTAATTCTCCAAGC
AACAGTAATGCCAATTGTGAATTTTCATTGAAAAACATCAGTGACATTCATCCAGAATCCTTCAA
CTTGCAAGTGGTTTTTCAATGACTCTTGTGCACCTCCTACTGGTTTCAATAAACCCACCAGTAACG
ACACCATGGCGAGTGGCTGGAGAGCATCTAGTTTCCACTTCGATTCTGAAGAAACAAACATAGG
CTTATCCACTTCTCAGTATTTTtaggtctattgcttggttggaattctggaggtcctggttgggct
CAGTCAGATAGTCATCGGTTTCCTTGGCTGTCTGTGTGGAGTCTCTAAGCGAAGAAGTCAAATTG
TGTAGTTTAAATGGGAATAAAATGTAAGTATCAGTAGTTTGAAAAAAAAA

FIGURE 48

MTCCEGWTS CNGFSLLV LLLGVVLNAIPLIVSLVEEDQFSQNPISCFEWWFPGIIGAGLMAIPA
TTMSLTARKRACCNRTGMFLSSFFSVITVIGALYCM LISIQALLKGPLMCNSPSNSNANCEFSL
KNISDIHPESFNLQWFFNDSCAPPTGFNKPTSNDTMASGWRASSFHF DSEENKHRLIHFSVFLGL
LLVGILEVLFGLSQIVIGFLGCLCGVSKRRSQIV

Important features:

Transmembrane domains:

amino acids 10-31 (type II), 50-72, 87-110, 191-213

N-glycosylation sites.

amino acids 80-84, 132-136, 148-152, 163-167

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 223-227

N-myristoylation sites.

amino acids 22-28, 54-60, 83-89, 97-103, 216-222

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 207-218

TNFR/NGFR family cysteine-rich region protein.

amino acids 4-12

FIGURE 49

ATCCGTTCTCTGCGCTGCCAGCTCAGGTGAGCCCTCGCCAAGGTGACCTCGCAGGACACTGGTGA
AGGAGCAGTGAGGAACCTGCAGAGTCACACAGTTGCTGACCAATTGAGCTGTGAGCCTGGAGCAG
ATCCGTGGGCTGCAGACCCCCGCCCAAGTGCCTCTCCCCCTGCAGCCCTGCCCCCTCGAACTGTGA
CATGGGAGAGAGTGACCCTGGCCCTTCTCCTACTGGCAGGCCTGACTGCCTTGGAAGCCAATGACC
CATTTGCCAATAAAGACGATCCCTTCTACTATGACTGGAAAAACCTGCAGCTGAGCGGACTGATC
TGCGGAGGGCTCCTGGCCATTGCTGGGATCGCGGCAGTTCTGAGTGGCAAATGCAAATACAAGAG
CAGCCAGAAGCAGCACAGTCCTGTACCTGAGAAGGCCATCCCACTCATCACTCCAGGCTCTGCCA
CTACTTGCTTGAGCACAGGACTGGCCTCCAGGGATGGCCTGAAGCCTAACACTGGCCCCCAGCACC
TCCTCCCCCTGGGAGGCCTTATCCTCAAGGAAGGACTTCTCTCCAAGGGCAGGCTGTTAGGCCCT
TTCTGATCAGGAGGCTTCTTTATGAATTAACTCGCCCCACCACCCCTCA

FIGURE 50

MERVTLALLLLAGLTALEANDPFANKDDPFYYDWKNLQLSGLICGGLLAIAAGIAAVLSGKCKYKS
SQKQHSPVPEKAIPILITPGSATTC

Important features:

Signal peptide:

amino acids 1-16

Transmembrane domain:

amino acids 36-59

N-myristoylation sites.

amino acids 41-47, 45-51, 84-90

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7.

amino acids 54-67

FIGURE 51

GTGGACTCTGAGAAGCCAGGCAGTTGAGGACAGGAGAGAGAAGGCTGCAGACCCAGAGGGAGGG
AGGACAGGGAGTCGGAAGGAGGAGGACAGAGGAGGGCACAGAGACGCAGAGCAAGGGCGGCAAGG
AGGAGACCTTGGTGGGAGGAAGACACTCTGGAGAGAGAGGGGGCTGGGCAGAGATGAAGTTCCAG
GGGCCCCCTGGCCTGCCTCCTGCTGGCCCTCTGCCTGGGCAGTGGGAGGCTGGCCCCCTGCAGAG
CGGAGAGGAAAGCACTGGGACAAATATTGGGGAGGCCCTTGGACATGGCCTGGGAGACGCCCTGA
GCGAAGGGGTGGGAAAGGCCATTGGCAAAGAGGCCGGAGGGGCAGCTGGCTCTAAAGTCAGTGAG
GCCCTTGGCCAAGGGACCAGAGAAGCAGTTGGCACTGGAGTCAGGCAGGTTCCAGGCTTTGGCGC
AGCAGATGCTTTGGGCAACAGGGTCGGGGAAGCAGCCCATGCTCTGGGAAACACTGGGCACGAGA
TTGGCAGACAGGCAGAAGATGTCTATTCGACACGGAGCAGATGCTGTCCGCGGCTCCTGGCAGGGG
GTGCCTGGCCACAGTGGTGCTTGGGAACTTCTGGAGGCCATGGCATCTTTGGCTCTCAAGGTGG
CCTTGGAGGCCAGGGCCAGGGCAATCCTGGAGGTCTGGGGACTCCGTGGGTCCACGGATACCCCG
GAAACTCAGCAGGCAGCTTTGGAATGAATCCTCAGGGAGCTCCCTGGGGTCAAGGAGGCAATGGA
GGGCCACCAAACCTTTGGGACCAACACTCAGGGAGCTGTGGCCAGCCTGGCTATGGTTCAAGTGA
AGCCAGCAACCAGAATGAAGGCTGCACGAATCCCCACCATCTGGCTCAGGTGGAGGCTCCAGCA
ACTCTGGGGGAGGCAGCGGCTCACAGTCGGGCAGCAGTGGCAGTGGCAGCAATGGTGACAACAAC
AATGGCAGCAGCAGTGGTGGCAGCAGCAGTGGCAGCAGCAGTGGCAGCAGCAGTGGCGGCAGCAG
TGGCGGCAGCAGTGGTGGCAGCAGTGGCAACAGTGGTGGCAGCAGAGGTGACAGCGGCAGTGAGT
CCTCCTGGGGATCCAGCACCGGCTCCTCCTCCGGCAACCACGGTGGGAGCGGCGGAGGAAATGGA
CATAAACC CGGTGTGAAAAGCCAGGGAATGAAGCCCGCGGGAGCGGGGAATCTGGGATT CAGGG
CTTCAGAGGACAGGGAGTTTCCAGCAACATGAGGGAAATAAGCAAAGAGGGCAATCGCCTCCTTG
GAGGCTCTGGAGACAATTATCGGGGGCAAGGGTCGAGCTGGGGCAGTGGAGGAGGTGACGCTGTT
GGTGGAGTCAATACTGTGAACTCTGAGACGTCTCCTGGGATGTTTAACTTTGACACTTCTTGAA
GAATTTTAAATCCAAGCTGGGTTTCATCAACTGGGATGCCATAAACAAGGACCAGAGAAGCTCTC
GCATCCCGTGACCTCCAGACAAGGAGCCACCAGATTGGATGGGAGCCCCCACTCCCTCCTTAA
AACACCACCCTCTCATCACTAATCTCAGCCCTTGCCCTTGAAATAAACCTTAGCTGCCCCACAAA
AA
AA

Year	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100
1970	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100	

Signal peptide:

amino acids 1-21

N-glycosylation site.

amino acids 265-269

Glycosaminoglycan attachment site.

amino acids 235-239, 237-241, 244-248, 255-259, 324-328, 388-392

Casein kinase II phosphorylation site.

amino acids 26-30, 109-113, 259-263, 300-304, 304-308

N-myristoylation site.

amino acids 17-23, 32-38, 42-48, 50-56, 60-66, 61-67, 64-70, 74-80,
90-96, 96-102, 130-136, 140-146, 149-155, 152-158, 155-161,
159-165, 163-169, 178-184, 190-196, 194-200, 199-205, 218-224,
236-242, 238-244, 239-245, 240-246, 245-251, 246-252, 249-252,
253-259, 256-262, 266-272, 270-276, 271-277, 275-281, 279-285,
283-289, 284-290, 287-293, 288-294, 291-297, 292-298, 295-301,
298-304, 305-311, 311-317, 315-321, 319-325, 322-328, 323-329,
325-331, 343-349, 354-360, 356-362, 374-380, 381-387, 383-389,
387-393, 389-395, 395-401

Cell attachment sequence.

amino acids 301-304

FIGURE 53

GGAGAAGAGGTTGTGTGGGACAAGCTGCTCCCGACAGAAGGATGTCGCTGCTGAGCCTGCCCTGG
CTGGGCCTCAGACCGGTGGCAATGTCCCATGGCTACTCCTGCTGCTGGTTGTGGGCTCCTGGCT
ACTCGCCCGCATCCTGGCTTGACCTATGCCTTCTATAACAACTGCCCGCGCTCCAGTGTTC
CACAGCCCCAAAACGGAACGGTTTTGGGGTCACCTGGGCCTGATCACTCCTACAGAGGAGGGC
TTGAAGGACTCGACCCAGATGTCGGCCACCTATCCCAGGGCTTTACGGTATGGCTGGGTCCCAT
CATCCCCCTTCATCGTTTTATGCCACCCTGACACCATCCGGTCTATCACCAATGCCTCAGCTGCCA
TTGCACCCAAGGATAATCTCTTCATCAGGTTCTTGAAGCCCTGGCTGGGAGAAGGGATACTGCTG
AGTGGCGGTGACAAGTGGAGCCGCCACCGTCGGATGCTGACGCCCGCCTTCCATTTCAACATCCT
GAAGTCCTATATAACGATCTTCAACAAGAGTGCAAACATCATGCTTGACAAGTGGCAGCACCTGG
CCTCAGAGGGCAGCAGTCGCTGGACATGTTTGGAGCACATCAGCCTCATGACCTTGACAGTCTA
CAGAAATGCATCTTCAGCTTTGACAGCCATTGTGAGGAGAGGCCAGTGAATATATTGCCACCAT
CTTGGAGCTCAGTGCCCTTGTAGAGAAAAGAAGCCAGCATATCCTCCAGCACATGGACTTTCTGT
ATTACCTCTCCCATGACGGGCGGCGCTTCCACAGGGCCTGCCGCTGGTGCATGACTTCACAGAC
GCTGTTCATCCGGGAGCGGCGTCGCACCCTCCCCACTCAGGGTATTGATGATTTTTTCAAAGACAA
AGCCAAGTCCAAGACTTTGGATTTTATTGATGTGCTTCTGCTGAGCAAGGATGAAGATGGGAAGG
CATTGTTCAGATGAGGATATAAGAGCAGAGGCTGACACCTTCATGTTTGGAGGCCATGACACCACG
GCCAGTGGCCTCTCCTGGGTCTGTACAACCTTGCGAGGCACCCAGAATACCAGGAGCGCTGCCG
ACAGGAGGTGCAAGAGCTTCTGAAGGACCGCGATCCTAAAGAGATTGAATGGGACGACCTGGCCC
AGCTGCCCTTCTGACCATGTGCGTGAAGGAGAGCCTGAGGTTACATCCCCCAGCTCCCTTCATC
TCCCGATGCTGCACCCAGGACATTGTTCTCCAGATGGCCGAGTCATCCCCAAAGGCATTACCTG
CCTCATCGATATTATAGGGGTCCATCACAACCCAACTGTGTGGCCGGATCCTGAGGTCTACGACC
CCTTCCGCTTTGACCCAGAGAACAGCAAGGGGAGGTACCTCTGGCTTTTATTCCCTTCTCCGCA
GGGCCCAGGAAGTGCATCGGGCAGGCGTTGCCATGGCGGAGATGAAAGTGGTCCCTGGCGTTGAT
GCTGCTGCACTTCCGGTTCCCTGCCAGACCACACTGAGCCCCGAGGAAGCTGGAATTGATCATGC
GCGCCGAGGGCGGGCTTTGGCTGCGGGTGAGCCCCCTGAATGTAGGCTTGCACTGACTTTCTGAC
CCATCCACCTGTTTTTTTGCAGATTGTCATGAATAAACGGTGCTGTCAA

FIGURE 54

MSLLSLPWLGLRPVAMSPWLLLLLVVGSWLLARILAWTYAFYNNCRRLQCFPPKRNWFWGHLG
LITPTEEGLKDSTQMSATYSQGFTVWLGPIIPFIVLCHPDTIRSITNASAAIAPKDNLFIRFLKP
WLGEGILLSGGDKWSRHRMLTPAFHFNILKSYITIFNKSANIMLDKWQHLASEGSSRLDMFEHI
SLMTLDSLQKCI FSFDSHCQERPSEYIATILELSALVEKRSQHILQHMDFLYYLSHDGRRFHRAC
RLVHDFTDVIRERRRTLPTQGIDDFKDKAKSKTLDFIDVLLLSKDEDGKALSDEDIRAEADTF
MFGGHDTTASGLSWVLYNLARHPEYQERCQEVQELLKDRDPKEIEWDDLAQLPFLTMCVKESLR
LHPPAPFISRCCTQDIVLPDGRVIPKGITCLIDIIGVHHNPTVWPDPEVYDPFRFDPENSKGRSP
LAFIPFSAGPRNCIGQAFAMAEMKVVLALMLLHFRFLPDHTEPRRKLELIMRAEGGLWLRVEPLN
VGLQ

Important features:

Transmembrane domains:

amino acids 13-32 (type II), 77-102

Cytochrome P450 cysteine heme-iron ligand signature.

amino acids 461-471

N-glycosylation sites.

amino acids 112-116, 168-172

FIGURE 55

ATCGCATCAATTGGGAGTACCATCTTCCTCATGGGACCAGTGAAACAGCTGAAGCGAATGTTTGA
GCCTACTCGTTTGATTGCAACTATCATGGTGCTGTTGTGTTTTGCACTTACCCTGTGTTCTGCCT
TTTGGTGGCATAACAAGGGACTTGCACTTATCTTCTGCATTTTGCAGTCTTTGGCATTGACGTGG
TACAGCCTTTCCTTCATACCATTTGCAAGGGATGCTGTGAAGAAGTGTGTTTGGCGTGTGCTTGC
ATAATTCATGGCCAGTTTTATGAAGCTTTGGAAGGCACTATGGACAGAAGCTGGTGGACAGTTTT
GTAACATCTTCGAAACCTCTGTCTTACAGACATGTGCCTTTTATCTTGCAGCAATGTGTGCTT
GTGATTCGAACATTTGAGGGTTACTTTTGGGAAGCAACAATACATTCTCGAACCTGAATGTCAGTA
GCACAGGATGAGAAGTGGGTCTGTATCTTGTGGAGTGAATCTTCCTCATGTACCTGTTTCCTC
TCTGGATGTTGTCCCACTGAATTCCCATGAATACAAACCTATTACAGCAACAGCAAAAAAAAAAAAA
AA

FIGURE 56

MGPVKQLKRMFEPTRLIATIMVLLCFALTLCSAFWWHNKGLALIFCILQSLALTWYSLSFIPFAR
DAVKKCFVCLA

Important features:

Signal peptide:

amino acids 1-33

Type II fibronectin collagen-binding domain protein.

amino acids 30-72

10063551

FIGURE 57

CGGCTCGAGCTCGAGCCGAATCGGCTCGAGGGGCGAGTGGAGCACCAGCAGGCGGCCAATGCTCTGTCTGTGCCTG
TACGTGCCGGTCAATCGGGGAAGCCAGACCGAGTTCAGTACTTTGAGTCGAAGGGGCTCCCTGCCGAGCTGAAGTCC
ATTTTCAAGCTCAGTGTCTTATCCCTCCAGGAATCTCCACCTACCGCCAGTGGAGCAGAAAAATGTACAAGCT
GGAGATAAGGACCTTGATGGGCGAGCTAGACTTTGAAGAATTTGTCCATTATCTCCAAGATCATGAGAAGAAGCTGAGG
CTGGTGTAAAGATTTTGACAAAAAGAAATGATGGACGCATTGACGCGCAGGAGATCATGCAGTCCCTGCCGGACTTG
GGAGTCAAGATATCTGAACAGCAGGCAGAAAAAATCTCAAGAGCATGGATAAAAACGGCAGCATGACCATCGACTGG
AACGAGTGGAGAGACTACCACCTCTCCACCCCGTGGAAAAACATCCCCGAGATCATCTCTACTGGAGCATTCACG
ATCTTTGATGTGGGTGAGAATCTAACGGTCCCGGATGAGTTCACAGTGGAGGAGAGGCAGACGGGGATGTGGTGGAGA
CACCTGGTGGCAGGAGGTGGGGCAGGGGCGTATCCAGAACCTGCACGGCCCCCTGGACAGGCTCAAGGTGCTCATG
CAGGTCATGCCTCCCGCAGCAACAACATGGGCATCGTTGGTGGCTTCACTCAGATGATTTCGAGAAGGAGGGGCCAGG
TCACTCTGGCGGGGCAATGGCATCAACGTCCTCAAAATGGCCCCGAATCAGGCATCAAAATCATGGCTATGAGCAG
ATCAAGCGCCTGTGTGGTAGTGACCAAGGAGACTCTGAGGATTCACGAGAGGCTTGTGGCAGGGTCCCTGGCAGGGGCC
ATGCCCCAGAGCAGCATCTACCAATGGAGGTCTGAAGACCCGATGGCGCTGGGAAGACAGGCCAGTACTCAGGA
ATGCTGGACTGCGCCAGGAGGATCTGGCCAGAGAGGGGTGGCCGCCTTCTACAAGGCTATGTCCCCAATCATGCTG
GGCATCATCCCTATGCCGGCATCGACCTTGCACTCTACGAGACGCTCAAGAATGCCTGGCTGCAGCACTATGCAGTG
AACAGCGCGGACCCCGGCTGTGTGTCTCTGGCCTGTGGCACCATGTCCAGTACCTGTGGCCAGCTGGCCAGCTAC
CCCCGGCCCTAGTCAGGACCCGGATGCAGGCGCAAGCCTCTATTGAGGGCGCTCCGGAGGTGACCATGAGCAGCCTC
TTCAAAACATATCTGCGGACCCGAGGGGGCTTCGGGCTGTACAGGGGGCTGGCCCCAACTTCATGAAGGTATCCCA
GCTGTGAGCATCAGTACCTGTCTACGAGAACCTGAAGATCACCTGGCGCTGCAGTCCGCGTGAAGGGGGAGGGG
CGCCCGGCGAGTGGACTCGCTGATCTGGGCGCAGCCTGGGGTGTGACGCCATCTCATTCTGTGAATGTGCCAACT
AAGCTGTCTCGAGCCAAGCTGTGAACCCCTAGACGCCACCCGAGGGAGGGTGGGGAGAGCTGGCAGGCCAGGGCTT
GTCTGTGACCCCCAGCAGACCTCTGTGTGGTTCAGCGAAGACCACAGGCATTCCTTAGGGTCCAGGGTCAAGAGG
CTCCGGGCTCACATGTGTAAAGCAGGACATTTCTGCAGTGCCTGCCAATAGTGAGCTTGGAGCTGGAGGCCGCT
TAGTCTCTCCATTTACCCCTTGACGCCAGCTGTGGCCACGGCCCTGCCCTCTGGTCTGCCGTGCATCTCCCTGTGC
CCTCTGTCTGCCTGTCTGTCTGAGGTAAAGTGGGAGGAGGGCTACAGCCACATCCACCCCTCTGCCAATCCC
ATAATCCATGATGAAGGTGAGGTACGCTGGCCTCCAGGCCTGACTTCCCAACCTACAGCATTCAGCCAACTTGGC
CTTGGGAGTGCAGGGGGCTCGGGCTGCCTGGCTGGCTGCACAGAAGGCAAGTGTGGGGCTCATGGTGTCTGAGCT
GGCTGGACCTGTGAGGATGGGCCCCACCTCAGAACCAAACTCACTGTCCCACTGTGGCATGAGGGCAGTGGAGCA
CCATGTTTGAAGGCGAAGGGCAGAGCGTTTGTGTCTGGGGAGGGAAGGAAAGGTGTTGGAGGCCCTTAATTATGG
ACTGTGGGAAAGGGTTTTGTCCAGAAGGACAAGCCGGACAAATGAGCGACTTCTGTGCTTCCAGAGGAAGACGAGG
GAGCAGGAGCTTGGCTGACTGCTCAGAGTCTGTTCTGACGCCCTGGGGTTCCTGTCCAACCCAGCAGGGGCGCAGC
GGGACAGCCCCACATTCACCTTGTGTCTGCTGTTGGAACCTATTTATTTTGTATTTTGAACAGAGTTATGTCTT
AACTATTTTTATAGATTTGTTAATTAATAGCTTGTCAATTTCAAGTTCAATTTTATTTATCATATTTATGTTATGGTT
GATTGTACCTTCCCAAGCCCGCCAGTGGGATGGGAGGAGGAGAGAAGGGGGCTTGGGCGCTGCAGTCACTCT
GTCCAGAGAAATTCCTTTTGGGACTGGAGGCAGAAAAAGCGGCCAGAGGCAGCAGCCCTGGCTCCTTTCTTTGGCAG
GTTGGGGAAGGGCTTCCCCCAGCCTTAGGATTTAGGGTTTGAAGTGGGGCGTGGAGAGAGAGGAGGAACCTCAAT
AACCTTGAAGGTGGAATCCAGTTATTTCTGCGCTGCGAGGGTTCTTTATTTTCACTCTTTTCTGAATGTCAAGGCAG
TGAGGTGCCTCTCACTGTGAATTTGTGGTGGGCGGGGCTGGAGGAGAGGGTGGGGGCTGGCTCCGTCCCTCCCAGC
CTTCTGCTGCCCTTGCTTAAACAATGCCGGCAACTGGCGACCTCACGGTTGCACTTCCATTCCACCAGAATGACCTGA
TGAGGAAATCTCAATAGGATGCAAGATCAATGCAAAATTTTATATATGAACATATAACTGGAGTCGTCAAAAG
CAAATTAAGAAAGATTGGACGTTAGAAGTTGTCATTTAAAGCAGCCTTCAATAAAGTTGTTTCAAGCTGAAAAA
AA

FIGURE 58

MLCLCLYVPVIGEAQTEFQYFESKGLPAELKSIFKLSVFIPSQEFSTYRQWKQKIVQAGDKDLDG
QLDFEETFVHYLQDHEKKRLRVFKILDKKNDGRIDAQEI MQSLRDLGVKISEQQAEKILKSMDKNG
TMTIDWNEWRDYHLLHPVENIPEIILYWKHSTIFDVGENLTVPDEFTVEERQTGMWWRHLVAGGG
AGAVSRTCTAPLDRCLKVLMQVHASRSNNMGIVGGFTQMIREGGARSLWRGNGINVLKIAPESAIK
FMAYEQIKRLVGSDQETLRIHERLVAGSLAGAIAQSSIYPMEVLKTRMALRK TGQYSGMLDCARR
ILAREGVAAFYKGYVPNMLGIIPYAGIDLAVYETLKNAWLQHYAVNSADPGVFVLLACGTMSSTC
GQLASYPLALVRTRMQAASIEGAPEVTMSSLFKHILRTEGAFGLYRGLAPNFMKVIPAVSISYV
VYENLKITLGVQSR

Important features:

Signal peptide:

amino acids 1-16

Putative transmembrane domains:

amino acids 284-304, 339-360, 376-394

Mitochondrial energy transfer proteins signature.

amino acids 206-215, 300-309

N-glycosylation sites.

amino acids 129-133, 169-173

Elongation Factor-hand calcium-binding protein.

amino acids 54-73, 85-104, 121-140

FIGURE 59

GGAAGGCAGCGGCAGCTCCACTCAGCCAGTACCCAGATACGCTGGGAACCTTCCCCAGCCATGGC
TTCCCTGGGGCAGATCCTCTTCTGGAGCATAATTAGCATCATCATTATTCTGGCTGGAGCAATTG
CACTCATCATTGGCTTTGGTATTTAGGGAGACACTCCATCACAGTCACTACTGTGCGCTCAGCT
GGGAACATTGGGGAGGATGGAATCCTGAGCTGCACCTTTTGAACCTGACATCAAACCTTCTGATAT
CGTGATACAATGGCTGAAGGAAGGTGTTTTAGGCTTGGTCCATGAGTTCAAAGAAGGCAAAGATG
AGCTGTGCGAGCAGGATGAAATGTTGAGAGCCGGACAGCAGTGTGCTGATCAAGTGATAGTT
GGCAATGCCTCTTTGCGGCTGAAAAACGTGCAACTCACAGATGCTGGCACCTACAAATGTTATAT
CATCACTTCTAAAGGCAAGGGGAATGCTAACCTTGAGTATAAACTGGAGCCTTCAGCATGCCGG
AAGTGAATGTGGACTATAATGCCAGCTCAGAGACCTTGCGGTGTGAGGCTCCCCGATGGTTCCCC
CAGCCCACAGTGGTCTGGGCATCCCAAGTTGACCAGGGAGCCAACCTTCTCGGAAGTCTCCAATAC
CAGCTTTGAGCTGAACCTCTGAGAAATGTGACCATGAAGTTGTGTCTGTGCTCTACAATGTTACGA
TCAACAACACATACTCCTGTATGATTGAAAATGACATTGCCAAAGCAACAGGGGATATCAAAGTG
ACAGAAATCGGAGATCAAAAGGCGGAGTCACTACAGCTGCTAAACTCAAAGGCTTCTCTGTGTGT
CTCTTCTTTCTTTGCCATCAGCTGGGCACTTCTGCCTCTCAGCCCTTACCTGATGCTAAAAATAAT
GTGCCTTGGCCACAAAAAGCATGCAAAGTCATTGTTACAACAGGGATCTACAGAATATTTTAC
CACCAGATATGACCTAGTTTTATATTTCTGGGAGGAAATGAATTCATATCTAGAAGTCTGGAGTG
AGCAAACAAGAGCAAGAAACAAAAAGAAGCCAAAGCAGAAGGCTCCAATATGAACAAGATAAAT
CTATCTTCAAAGACATATTAGAAGTTGGGAAAATAATTCATGTGAAGTAGACAAGTGTGTTAAGA
GTGATAAGTAAATGCACGTGGAGACAAGTGCATCCCCAGATCTCAGGGACCTCCCCCTGCCTGT
CACCTGGGGAGTGAGAGGACAGGATAGTGCATGTTCTTTGTCTCTGAATTTTGTATATGTGC
TGTAATGTTGCTCTGAGGAAGCCCCTGGAAAGTCTATCCCAACATATCCACATCTTATATCCAC
AAATTAAGCTGTAGTATGTACCTAAGACGCTGCTAATTGACTGCCACTTCGCAACTCAGGGGCG
GCTGCATTTTAGTAATGGGTCAAATGATTCACTTTTTATGATGCTTCCAAAGGTGCCTTGGCTTC
TCTTCCCAACTGACAAATGCCAAAGTTGAGAAAAATGATCATAATTTTAGCATAAACAGAGCAGT
CGGGGACACCGATTTTATAAATAAACTGAGCACCTTCTTTTAAACAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 6o

MASLGQILFWSIISIIIIILAGAIALIIGFGISGRHSITVTTVASAGNIGEDGILSCTFEPDIKLS
DIVIQWLKEGVLGLVHEFKEGKDELSEQDEMFRGRTAVFADQVIVGNASRLKNVQLTDAGTYKC
YIITSKKGKGNANLEYKTGAFSMPEVNVVDYNASSETLRCEAPRWFPQPTVVWASQVDQGAFSEVS
NTSFELNSENVTMKVVSVLNVNTINNTYSCMIENDIAKATGDIKVTSEIKRRSHLQLLNSKASL
CVSSFFAISWALLPLSPYMLLK

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 258-281

N-glycosylation sites.

amino acids 112-116, 160-164, 190-194, 196-200, 205-209, 216-220,
220-224

N-myristoylation sites.

amino acids 52-58, 126-132, 188-194

FIGURE 61

TGACGTCAGAATCACCATGGCCAGCTATCCTTACCGGCAGGGCTGCCAGGAGCTGCAGGACAAG
CACCAGGAGCCCCCTCCGGGTAGCTACTACCCTGGACCCCCCAATAGTGGAGGGCAGTATGGTAGT
GGGCTACCCCCCTGGTGGTGGTTATGGGGGTCTGCCCTGGAGGGCCTTATGGACCACCAGCTGG
TGGAGGGCCCTATGGACACCCCAATCCTGGGATGTTCCCCTCTGGAACCTCAGGAGGACCATATG
GCGGTGCAGCTCCCGGGGGCCCCCTATGGTCAGCCACCTCCAAGTTCCTACGGTGCCAGCAGCCT
GGGCTTTATGGACAGGGTGGCGCCCCCTCCCAATGTGGATCCTGAGGCCTACTCCTGGTTCCAGTC
GGTGGACTCAGATCACAGTGGCTATATCTCCATGAAGGAGCTAAAGCAGGCCCTGGTCAACTGCA
ATTGGTCTTCATTCAATGATGAGACCTGCCTCATGATGATAAACATGTTTGACAAGACCAAGTCA
GGCCGCATCGATGTCTACGGCTTCTCAGCCCTGTGGAAATTCATCCAGCAGTGGAAGAACCTCTT
CCAGCAGTATGACCGGGACCGCTCGGGCTCCATTAGCTACACAGAGCTGCAGCAAGCTCTGTCCC
AAATGGGCTACAACCTGAGCCCCAGTTCACCCAGCTTCTGGTCTCCCGCTACTGCCACGCTCT
GCCAATCCTGCCATGCAGCTTGACCGCTTCATCCAGGTGTGCACCCAGCTGCAGGTGTGACAGA
GGCCTTCCGGGAGAAGGACACAGCTGTACAAGGCAACATCCGGCTCAGCTTCGAGGACTTCGTCA
CCATGACAGCTTCTCGGATGCTATTGACCCCAACCATCTGTGGAGAGTGGAGTGCACCAGGGACCTT
TCCTGGCTTCTTAGAGTGAGAGAAGTATGTGGACATCTCTTCTTTCTGTCCCTCTAGAAGAAC
ATTCTCCCTTGCTTGATGCAACACTGTTCCAAAAGAGGGTGGAGAGTCTGCATCATAGCCACCA
AATAGTGAGGACCGGGCTGAGGCCACACAGATAGGGGCCTGATGGAGGAGAGGATAGAAGTTGA
ATGTCTGATGGCCATGAGCAGTTGAGTGGCACAGCCTGGCACCAGGAGCAGGTCCTTGTAATGG
AGTTAGTGTCCAGTCAGCTGAGCTCCACCCTGATGCCAGTGGTGAGTGTTCATCGGCCTGTTACC
GTTAGTACCTGTGTTCCCTCACCAGGCCATCCTGTCAAACGAGCCCATTTTCTCCAAAGTGGAAT
CTGACCAAGCATGAGAGAGATCTGTCTATGGGACCAGTGGCTTGGATTCTGCCACACCCATAAAT
CCTTGTGTGTTAACTTCTAGCTGCCTGGGGCTGGCCCTGCTCAGACAAATCTGCTCCCTGGGCAT
CTTTGGCCAGGCTTCTGCCCCCTGCAGCTGGGACCCCTCACTTGCCCTGCCATGCTCTGCTCGGCT
TCAGTCTCCAGGAGACAGTGGTCACCTCTCCCTGCCAATACTTTTTTAATTGCAATTTTTTTTC
ATTTGGGGCCAAAAGTCCAGTGAAATTGTAAGCTTCAATAAAAGGATGAAACTCTGA

FIGURE 62

MASYPYRQGPCGAAGQAPGAPPGSYYPGPPNSGGQYGSGLPPGGGYGGPAPGGPYGPPAGGGPYG
HPNPGMFPSGTPGGPYGGAAPGGPYGQPPSSYGAQQPGLYGQGGAPPNVDPEAYSWFQSVDS DH
SGYISMKELKQALVNCNWSSFNDETCLMMINMFDKTKSGRIDVYGFSALWKFIQQWKNLFQQYDR
DRSGSISYTELQQALSQMGYNLSPQFTQLLVSRYCPRSANPAMQLDRFIQVCTQLQVLTEAFREK
DTAVQGNIRLSFEDFVTMTASRML

Important features of the protein:

Signal peptide:

amino acids 1-19

N-glycosylation site.

amino acids 147-150

Casein kinase II phosphorylation sites.

amino acids 135-138, 150-153, 202-205, 271-274

N-myristoylation sites.

amino acids 9-14, 15-20, 19-24, 33-38, 34-39, 39-44, 43-48, 61-
66, 70-75, 78-83, 83-88, 87-92, 110-115

FIGURE 63

CAGGATGCAGGGCCGCGTGGCAGGGAGCTGCGCTCCTCTGGGCCTGCTCCTGGTCTGTCTTCATC
TCCCAGGCCTCTTTGCCCGGAGCATCGGTGTTGTGGAGGAGAAAGTTTCCCAAACTTCGGGACC
AACTTGCCCTCAGCTCGGACAACCTTCCTCCACTGGCCCCCTCTAACTCTGAACATCCGCAGCCCGC
TCTGGACCCTAGGTCTAATGACTTGGCAAGGGTTCCTCTGAAGCTCAGCGTGCCTCCATCAGATG
GCTTCCCACCTGCAGGAGGTTCTGCAGTGCAGAGGTGGCCTCCATCGTGGGGGCTGCCTGCCATG
GATTCTGGCCCCCTGAGGATCCTTGGCAGATGATGGCTGCTGCGGCTGAGGACCGCCTGGGGGA
AGCGCTGCCTGAAGAACTCTCTTACCTCTCCAGTGCTGCGGCCCTCGCTCCGGGCAGTGGCCCTT
TGCCTGGGGAGTCTTCTCCCGATGCCACAGGCCTCTCACCTGAGGCTTCACTCCTCCACCAGGAC
TCGGAGTCCAGACGACTGCCCCGTTCTAATTCAGTGGGAGCCGGGGGAAAAATCCTTTCCCAACG
CCCTCCCTGGTCTCTCATCCACAGGGTTCTGCCGTGATCACCCCTGGGGTACCCTGAATCCAGTG
TGTCTGGGGAGGTGGAGGCCCTGGGACTGGTTGGGGAACGAGGCCCATGCCACCCCTGAGGGA
ATCTGGGGTATCAATAATCAACCCCCAGGTACCAGCTGGGGAAATATTAATCGGTATCCAGGAGG
CAGCTGGGGAAATATTAATCGGTATCCAGGAGGCAGCTGGGGGAATATTAATCGGTATCCAGGAG
GCAGCTGGGGGAATATTCATCTATAACCCAGGTATCAATAACCCATTTCTCCTGGAGTTCTCCGC
CCTCCTGGCTCTTCTTGGAACATCCCAGCTGGCTTCCCTAATCCTCCAAGCCCTAGGTTGCAGTG
GGGCTAGAGCACGATAGAGGGAAACCCAACATTGGGAGTTAGAGTCTGCTCCCGCCCCCTTGCTG
TGTGGGCTCAATCCAGGCCCTGTAAACATGTTTCCAGCACTATCCCCACTTTTCAGTGCCTCCCC
TGCTCATCTCCAATAAAATAAAAGCACTTATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAA

FIGURE 64

MQGRVAGSCAPLGLLLVLCLHLPGLFARSIGVVEEKVSNFGTNLPQLGQPSSTGPSNSEHPQPAL
DPRSNDLARVPLKLSVPPSDGFPPAGGSAVQRWPPSWGLPAMDSWPPEDPWQMAAAAEDRLGEA
LPEELSYLSSAAALAPGSGPLPGESSPDATGLSPEASLLHQDSESRRLPRSNSLGAGGKILSQRP
PWSLIHRVLPDHPWGTLNPSVSWGGGGPGTGWGTRPMPHPEGIWGINNQPPGTSGWNINRYPGGS
WGNINRYPGGSWGNINRYPGGSWGNIHLYPGINNPFPPGVLRPPGSSWNIPAGFPNPPSPRLQWG

Important features of the protein:

Signal peptide:

amino acids 1-26

Casein kinase II phosphorylation sites.

amino acids 56-59, 155-158

N-myristoylation sites.

amino acids 48-53, 220-225, 221-226, 224-229, 247-252, 258-263,
259-264, 269-274, 270-275, 280-285, 281-286, 305-310

FIGURE 65

AAGGAGAGGCCACCGGACTTCAGTGTCTCCTCCATCCCAGGAGCGCAGTGGCCACTATGGGGTC
TGGGCTGCCCCCTTGTCCTCCTCTTGACCCTCCTTGGCAGCTCACATGGAACAGGGCCGGGTATGA
CTTTGCAACTGAAGCTGAAGGAGTCTTTTCTGACAAATTCCTCCTATGAGTCCAGCTTCCTGGAA
TTGCTTGAAAAGCTCTGCCTCCTCCTCCATCTCCCTTCAGGGACCAGCGTCACCCCTCCACCATGC
AAGATCTCAACACCATGTTGTCTGCAACACATGACAGCCATTGAAGCCTGTGTCTTCTTGCCCC
GGGCTTTTGGGCCGGGGATGCAGGAGGCAGGCCCCGACCCTGTCTTTCAGCAGGCCCCACCCCTC
CTGAGTGGCAATAAATAAAATTCCGGTATGCTG

FIGURE 66

MGSGLPVLVLLLTLLGSSHGTGPGMTLQLKLKESFLTNSSEYESSFLELLEKLCLLLHLPSTSVTL
HHARSQHHVVCNT

Important features:

Signal peptide:

amino acids 1-19

N-glycosylation site.

amino acids 37-41

N-myristoylation sites.

amino acids 15-21, 19-25, 60-66

FIGURE 67

ACGGACCGAGGGTTCGAGGGAGGGACACGGACCAGGAACCTGAGCTAGGTCAAAGACGCCCGGGC
CAGGTGCCCCGTGCGAGGTGCCCCCTGGCCGGAGATGCGGTAGGAGGGGCGAGCGCGAGAAGCCCC
TTCCTCGGCGCTGCCAACCCGCCACCCAGCCCATGGCGAACCCCGGGCTGGGGCTGCTTCTGGCG
CTGGGCCTGCCGTTCTGTGCTGGCCCGCTGGGGCCGAGCCTGGGGGCAAATACAGACCACTTCTGC
AAATGAGAATAGCACTGTTTTGCCTTCATCCACCAGCTCCAGCTCCGATGGCAACCTGCGTCCGG
AAGCCATCACTGCTATCATCGTGGTCTTCTCCCTCTTGGCTGCCTTGCTCCTGGCTGTGGGGCTG
GCACTGTTGGTGCGGAAGCTTCGGGAGAAGCGGCAGACGGAGGGCACCTACCGGCCCAGTAGCGA
GGAGCAGTTCTCCCATGCGCCGAGGCCCGGGCCCCCTCAGGACTCCAAGGAGACGGTGCAGGGCT
GCCTGCCCATCTAGGTCCCCTCTCCTGCATCTGTCTCCCTTCATTGCTGTGTGACCTTGGGGAAA
GGCAGTGCCCTCTCTGGGCAGTCAGATCCACCCAGTGCTTAATAGCAGGGAAGAAGGTACTTCAA
AGACTCTGCCCCCTGAGGTCAAGAGAGGATGGGGCTATTCACTTTTATATATTTATATAAAATTAG
TAGTGAGATGTAAAAAAAAAAAAAAAAAAAA

FIGURE 68

MANPGLGLLLALGLPFLLARWGRAWGQIQTTSANENSTVLPSTSSSSDGNLRPEAITAIIVVFS
LLAALLLAVGLALLVRKLRQTEGTYRPSSEEQFSHAAEARAPQDSKETVQGCLPI

Important features:

Signal peptide:

amino acids 1-19

Transmembrane domain:

amino acids 56-80

N-glycosylation site.

amino acids 36-40

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 86-90

Tyrosine kinase phosphorylation site.

amino acids 86-94

N-myristoylation sites.

amino acids 7-13, 26-32

FIGURE 69

GCCAGGAATAACTAGAGAGGAACAATGGGGTTATTGAGAGGTTTTGTTTTCTCTTAGTTCTGTGCCTGCTGCACCAG
 TCAAATACTTCCTTCATTAAGCTGAATAATAATGGCTTTGAAGATATTGTCATTGTTATAGATCCTAGTGTGCCAGAA
 GATGAAAAAATAATTGAACAAATAGAGGATATGGTGACTACAGCTTCTACGTACCTGTTTGAAGCCACAGAAAAAGA
 TTTTTTTTCAAAAATGTATCTATATTAATTCCTGAGAATTGGAAGGAAAAATCCTCAGTACAAAAGGCCAAAAACATGAA
 AACCATAAACATGCTGATGTTATAGTTGCACCACCTACACTCCAGGTAGAGATGAACCATACACCAAGCAGTTCACA
 GAATGTGGAGAGAAAGGCGAATACATTCACTTCACCCCTGACCTTCTACTTGGAAAAAACAATAATGAATATGGACCA
 CCAGGCCAACTGTTTGTCCATGAGTGGGCTCACCTCCGGTGGGGAGTGTGATGAGTACAATGAAGATCAGCCTTTC
 TACCGTGCTAAGTCAAAAAAATCGAAGCAACAAGGTGTTCCGCAGGTATCTCTGGTAGAAATAGAGTTTATAAGTGT
 CAAGGAGGCAGCTGTCTTAGTAGAGCATGCAGAATTGATTCTACAAACAACTGTATGGAAAAAGATTGTCAATTCTTT
 CCTGATAAAGTACAACAGAAAAAGCATCCATAATGTTTATGCAAGTATTGATTCTGTTGTTGAATTTTGAACGAA
 AAAACCCATAATCAAGAAGCTCCAAGCCTACAAAACATAAAGTGAATTTTGAAGTACATGGGAGGTGATTAGCAAT
 TCTGAGGATTTTAAAAACACCATACCCATGGTGACACCACCTCCCTCCACCTGTCTCTCATGCTGAAGATCAGTCAA
 AGAATTGTGTGCTTAGTCTTGATAAGTCTGGAAGCATGGGGGTAAGGACCGCTAAATCGAATGAATCAAGCAGCA
 AAACATTTCTGCTGCAGACTGTTGAAAAATGGATCCTGGGTGGGGATGTTTCACTTTGATAGTACTGCCACTATTGTA
 AATAAGCTAATCCAATAAAAGCACTGATGAAAGAAACACACTCATGGCAGGATTACCTACATATCCTCTGGGAGGA
 ACTTCCATCTGCTCTGGAAATTAATATGCATTTGAGGTGATTGGAGAGCTACATTCCTCAACTCGATGGATCCGAAGTA
 CTGCTGCTGACTGATGGGGAGGATAACACTGCAAGTCTTGTATTGATGAAGTGAACAAAGTGGGGCCATTGTTTCTAT
 TTTATTGCTTTGGGAAGAGCTGCTGATGAAGCAGTAATAGAGATGAGCAAGATAACAGGAGGAAGTCATTTTATGTT
 TCAGATGAAGCTCAGAACATGGCCTCATTGATGCTTTTGGGGCTTACATCAGGAAATACTGATCTCTCCAGAG
 TCCCTTCAGCTCGAAAGTAAGGATTAACACTGAATAGTAATGCTGCGATGAACGACACTGTCAATAATTGATAGTACA
 GTGGGAAAGGACACGTTCTTTCTCATCACATGGAACAGTCTGCCCTCCAGTATTTCTCTCTGGGATCCCAAGTGAACA
 ATAATGGAAAAATTCACAGTGGATGCAACTTCCAAAATGGCCTATCTCAGTATTCAGGAACTGCAAAAGGTGGGCACT
 TGGGCATACAATCTTCAAGCCAAAGCGAACCCAGAAACATTAACATTTACAGTAACCTCTCGAGCAGCAAAATCTTCT
 GTGCCCTCAATCACAGTGAATGCTAAAATGAATAAGGACGTAACAGTTTCCCAGCCCAATGATTGTTTACGCGAGAA
 ATTCTACAAGGATATGTACCTGTTCTTGGAGCCAATGTGACTGCTTTTATTGAATCACAGAATGGACATACAGAAGTT
 TTGGAACTTTTGGATAATGGTGCAGGCGCTGATTCTTTCAAGAAATGATGGAGTCTACTCCAGGTATTTTACAGCATAT
 ACAGAAAATGGCAGATATAGCTTAAAGTTGCGGCTCATGGAGGAGCAACACTGCCAGGCTAAAATTACGGCCTCCA
 CTGAATAGAGCCGCTACATACCAGGCTGGGTAGTGAACGGGGAAATTGAAGCAAAACCCGCCAAGACCTGAAATTGAT
 GAGGATACTCAGACCACCTTGGAGGATTTGAGCCGAACAGCATCCGGAGGTGCATTTGTGGTATCACAGTCCCAAGC
 CTTCCTTGCCTGACCAATACCCACCAAGTCAAATCACAGACCTTGATGCCACAGTTTATGAGGATAAGATTATTCTT
 ACATGGACAGCACCAGGAGATAATTTTGTGTTGGAAAAGTTCAACGTTATATCATAAGAATAAGTGCAAGTATTCTT
 GATCTAAGAGACAGTTTGTGATGCTCTTCAAGTAATACTACTGATCTGTCCACAAAGGAGGCCAACTCCAAGGAA
 AGCTTTGCATTTAAACCAGAAAAATATCTCAGAGAAAAATGCAACCCACATATTATTGCCATTAAAAATATAGATAAA
 AGCAATTTGACATCAAAAGTATCCAACATTGCACAAGTAACTTTGTATCCCTCAAGCAAATCCTGATGACATTGAT
 CCTACACCTACTCCTACTCCTACTCCTACTCCTGATAAAAGTCATAATTCTGGAGTTAATATTCTACGCTGGTATTG
 TCTGTGATTGGGTCTGTTGTAATTGTTAACTTTATTTTAAAGTACCACCATTTGAACCTTAACGAAGAAAAAATCTTC
 AAGTAGACCTAGAAGAGACTTTTAAAAACAAAAAATGTAAGTAAAGGATATTCTGAATCTTAAAAATTCATCCCAT
 GTGTGATCATAAACTCATAAAAAATAATTTTAAAGTGTCCGAAAAGGATACTTTGATTAAATAAAACACTCATGGATA
 TGTAAAAACTGTCAAGATTAAAAATTAATAGTTTCATTTATTGTTATTTTGTAAAGAAATAGTGATGAACAAAG
 ATCCTTTTTCATACTGATACCTGGTTGTATATTATTGATGCAACAGTTTTCTGAAATGATATTTCAAATTGCATCAA
 GAAATTAATAATCATCTATCTGAGTAGTCAAAATACAAGTAAGGAGAGCAATAAACAACATTTGGAAAAAATAA
 AA

FIGURE 70

MGLFRGFVFLVLVCLLHQSNSTFIKLNNNGFEDIVIVIDPSVPEDEKIIIEQIEDMVTTASTYLFE
ATEKRFFFKNVSILIPENWKENPQYKRPKHENHKHADVIVAPPTLPGRDEPYTKQFTECGEKGEY
IHFTPDLLLGGKKQNEYGPPGKLFVHEWAHLRWGVFDEYNEDQPFYRAKSKKIEATRCISAGISGRN
RVYKCQGGSCLSRACRIDSTTKLYGKDCQFFPDQVTEKASIMFMQSIDSVVEFCNEKTHNQEAP
SLQNIKCNFRSTWEVISNSEDfKNTIPMVTPPPPVFSLLKISQRIVCLVLDKSGSMGGKDLNR
MNQAAKHFLQLTVENGSWVGMVHFDSTATIVNKLIQIKSSDERNTLMAGLPTYPLGGTSICSGIK
YAFQVIGELHSQLDGEVLLLLTDGEDNTASSCIDEVKQSGAIVHFIALGRAADEAVIEMSKITGG
SHEFYVSDEAQNNGLIDAFGALTSGNTDLSQKSLQLESKGLTLNSNAWMNDTVIIDSTVGKDTFFL
ITWNSLPPSISLWDPSGTIMENFTVDATSKMAYLSIPGTAKVGTWAYNLQAKANPETLTITVTSR
AANSSVPPITVNAKMNDVNSFPSPMIVYAEILQGYVPVLGANVTAFIESQNGHTEVLELLDNGA
GADSKNDGVYSRYFTAYTENGRYSLKVRHGGANTARLKLRLPLNRAAYIPGWVNGEIEANPP
RPEIDEDTQTTLDEFSTRASGGAFVVSQVPSLPLPDQYPPSQITDLDATVHEDKIILTWTAPGDN
FDVGKVQRYIIRISASILDLRDSFDDALQVNTTDLSPKEANSKESFAFKPENISEENATHIFIAI
KSIDKSNLTSKVSNIAQVTLFIPQANPDDIDPTPTPTPTPTPKSHNSGVNISTLVLSVIGSVVI
VNFILSTTI

Signal peptide:

amino acids 1-21

Putative transmembrane domains:

amino acids 284-300, 617-633

Leucine zipper pattern.

amino acids 469-491, 476-498

N-glycosylation site.

amino acids 20-24, 75-79, 340-344, 504-508, 542-546, 588-592,
628-632, 811-815, 832-836, 837-841, 852-856, 896-900

FIGURE 71

CTCCTTAGGTGGAACCCCTGGGAGTAGAGTACTGACAGCAAAGACCGGGAAAGACCATACGTCCCCGGGCAGGGGTGA
CAACAGGTGTCTATCTTTTGTATCTCGTGTGTGGCTGCCTTCTCTATTCAAGGAAAGACGCCAAGGTAATTTTGACCCA
GAGGAGCAATGATGTAGCCACCTCCTAACCTTCCCTTCTTGAACCCCAAGTTATGCCAGGATTTACTAGAGAGTGTCA
ACTCAACCAGCAAGCGGCTCCTTCGGCTTAACTTGTGGTTGGAGGAGAGAACCTTTGTGGGGCTGCGTTCTCTTAGCA
GTGCTCAGAAGTGACTTGCTGAGGGTGGACCAGAAGAAAGGAAAGGTCCCCTCTTGTCTGTGGCTGCACATCAGGAA
GGCTGTGATGGGAATGAAGGTGAAAACCTTGGAGATTTCACTTCAGTCATTGCTTCTGCCTGCAAGATCATCTTTAAA
AGTAGAAGAGCTGCTCTGTGTGGTGGTTAACTCCAAGAGGCAGAACTCGTTCTAGAAGGAAATGGATGCAAGCAGCTC
CGGGGGCCCCAAACGCATGCTTCTGTGGTCTAGCCCAGGGAAGCCCTTCCGTGGGGGCCCGGCTTTGAGGGATGCC
ACCGGTTCTGGACGCATGGCTGATTCTGATGATGATGGTTCCGGCGGGGGCTGCTTGCCTGGATTTCCCGGGTGGTG
GTTTTGCTGGTGTCTCTGTGTGTGCTATCTGTGCTGTACATGTTGGCTGCACCCCAAGGTGACGAGGAGCAG
CTGGCACTGCCAGGGCCAAACAGCCCCACGGGGAAGGAGGGGTACCAGGCCGCTCTTCAGGAGTGGGAGGAGCAGCAC
CGCAACTACGTGAGCAGCCTGAAGCGGCAGATCGCACAGCTCAAGGAGGAGCTGCAGGAGAGGAGTGACGAGCTCAGG
AATGGGCAGTACCAAGCCAGCGATGCTGCTGGCTGGGTCTGGACAGGAGCCCCCAGAGAAAACCCAGGCCGACCTC
CTGGCCCTTCTGCACTCGCAGGTGGACAAGGCAGAGGTGAATGCTGGCGTCAAGCTGGCCACAGAGATGACAGCAGTG
CCTTTCGATAGCTTTACTCTACAGAAGGTGTACAGCTGGAGACTGGCCTTACCCGCCACCCGAGGAGAAGCCTGTG
AGGAAGGACAGCGGGATGAGTTGGTGGAGGCCATTGAATCAGCCTTGGAGACCCCTGAACAATCCTGCAGAGAACAGC
CCCAATCACCGTCCCTTACACGGCCTCTGATTTCTAGAGAGGATCTACCGAACAGAAAGGACAAAGGGACATTGTAT
GAGCTCACCTTCAAAGGGGACCACAAACAGCAATTCAAACGGCTCATCTTATTTGACCATTGAGCCCCATCATGAAA
GTGAAAAATGAAAAGCTCAACATGGCCCAACAGCTTATCAATGTTATCGTGCCTCTAGCAAAAAGGGTGGACAAGTTC
CGGCAGTTTATGCAATTTTCAAGGAGATGTGATTGAGCAGGATGGGAGAGTCCATCTCACTGTTGTTTACTTTGGG
AAAGAAAGAAATAAATGAAGTCAGGGAATACTTGAACACTTCCAAAGCTGCCAACTTCAGGAACCTTACCTTCATC
CAGCTGAATGGAGAATTTCTCGGGGAAAGGACTTGATGTTGGAGCCCGCTTCTGGAAGGGAAGCAAGCTCCTTCTC
TTTTTCTGTGATGTGGACATCTACTTCACATCTGAATTCCTCAATACGTGTAGGCTGAATACAGCCAGGGAAGAAG
GTATTTTATCCAGTTCTTTTTCAGTCAGTACAATCTGGCATAATATACGCCACCATGATGCAGTCCCTCCCTTGGAA
CAGCAGCTGGTCATAAAGAGGAACTGGATTTTGGAGAGACTTTGGATTGGGATGACGTGTGAGTATCGGTGAGAC
TTTATCAATATAGGTGGGTTTGTATCTGGACATCAAGGCTGGGGCGGAGAGGATGTGCACCTTTATCGCAAGTATCTC
CCCAATCACCGTCCCTTACACGGCCTCTGATTTCTAGAGAGGATCTTCCACCTCTGGCATGAGAAGCCTGTCATGGACGAG
CTGACCCCGGAGCAGTACAAGATGTGATGCACTCCAGGCCATGAACGAGGCATCCACGGCCAGCTGGGCATGCTG
GTGTTGAGGCACGAGATAGAGGCTCACCTTCGCAAAACAGAAACAGAAAGCAAGTAGCAAAAAACATGAAGTCCCAGA
GAAGGATTTGGGAGACACTTTTTCTTTCTTTTGAATTAAGTGGCTGCAACAGAGAAAGACTTCCATAAA
GGACGCAAAAAGAAATGGACTGATGGGTGAGAGATGAGAAAGCCTCCGATTTCTCTGTTGGGCTTTTACAACAGA
AATCAAATCTCCGCTTTGCCTGCAAAAGTAACCCAGTTGCACCTGTGAAGTGTCTGACAAAGGCAGAAATGCTTGTG
AGATTATAAGCCTAATGGTGTGGAGGTTTGTGATGGTGTTTACAATACACTGAGACCTGTTGTTTGTGTGCTCATTGA
AATATTTCATGATTTAAGAGCAGTTTGTAAAAAATTCATTAGCATGAAAGGCAAGCATATTTCTCCTCATATGAATGA
GCCTATCAGCAGGGCTCTAGTTTCTAGGAATGCTAAAAATATCAGAAGGCAGGAGAGATAGGCTTATTATGATACT
AGTGAGTACATTAAATGAAATGAAATGGACAGAAAAAGAAAGAAACCATAAATATCGTGTCTATATTTCCCAAGAT
TAACCAAAATATCTGCTTATCTTTTGGTTGTCCTTTTAACTGTCTCCGTTTTTTTCTTTTATTTAAAAATGCACT
TTTTTTCCCTTGTGAGTTATAGTCTGCTTATTTAATTACCACTTTGCAAGCCTTACAGAGAGCACAAGTTGGCCTAC
ATTTTATATTTTAAAGAGATCTTTGAGATGCATTATGAGAACTTTCAGTTCAAAGCATCAATTTGATGCCATAT
CCAAGGACATGCCAATGCTGATTTCTGTCAGGCCTGAATGTCAGGCATTGAGACATAGGGAAGGAATGGTTGTACT
AATACAGACGTACAGATACTTTCTCTGAAGAGTATTTTCGAAGAGGAGCAACTGAACACTGGAGGAAAAGAAAATGAC
ACTTCTGCTTTACAGAAAAGGAACTCATTGAGACTGGTGATATCGTGATGTACCTAAAAGTCAGAAAACCATTTT
CTCCTCAGAAGTAGGGACCGCTTCTTACCTGTTAAATAAACCAAAGTATACCGTGTGAACCAACATCTCTTTTC
AAAACAGGGTGTCTCCTCTGGCTTCTGGCTTCCATAAGAAGAAATGGAGAAAATATATATATATATATATATATTGT
GAAAGATCAATCCATCTGCCAGAACTAGTGGGATGGAAGTTTTTGTACATGTTATCCACCCAGGCCAGGTGGAAG
TAAGTGAATTTTAAATTAAGCAGTTCTACTCAATCACCAGATGCTTCTGAAAATGCAATTTTATTACCATTT
CAAATATTTTAAAAATAAATACAGTTAAATAGAGTGGTTTCTTCAATCATGTGAAAATATTAGCCAGCACCAG
ATGCATGAGCTAATTTATCTCTTGTGCTTCTGCTTCTGTTTGTGCTCACAGTAACTCATTGTTTAAAGCTTCAAGAAC
ATTCAAGCTGTTGGTGTGTTAAAAATGCATTGTATTGATTGTACTGTTAGTTTATGAAATTTAATTTAAACACAGG
CCATGAATGGAAGGTGGTATTGCACAGCTAATAAAATATGATTGTGGATATGAA

FIGURE 72

MMVRRGLLAWISRVVLLVLLCCAISVLYMLACTPKGDEEQLALPRANSPGKEGYQAVLQEW
EQHRNYVSSLKRQIAQLKEELQERSEQLRNGQYQASDAAGLGDRSPPEKTQADLLAFLHSQV
AEVNAGVKLATEYAAVPPDSFTLQKVYQLETGLTRHPEEKPVKDKRDELVEAIESALET
ENSPNHRPYTASDFIEGIYRTERDKGTLYELTFKGDHKHEFKRLILFRPFSPIMKV
NEKLNMAN TLINVIPLAKRVDKFRQFMQNFREMCIEQDGRVHLTVVYFGKEEINEV
KGILENTSKAANFRNF TFIQLNGEFSRGKGLDVGARFWKGSNVLLFFCDVDIYFTSE
FLNTCRLNTQPGKKVFYPVLFSQY NPGIYGHHDVPPLEQQLVIKKETGEWRDFG
FGMTCQYRSDFINIGGFDDIKGWGGEDVHLYR KYLHSLNLIIVVRTPVRGLFHL
WHEKRCMDLTPQYKCMQSKAMNEASHGQLGMLVFRHEIEAHL RKQKQKTSSKKT

Important features:

Signal peptide:

amino acids 1-27

N-glycosylation sites.

amino acids 315-319, 324-328

N-myristoylation sites.

amino acids 96-102, 136-142, 212-218, 311-317, 339-345, 393-399

Amidation site.

amino acids 377-381

FIGURE 73

GAGACTGCAGAGGGAGATAAAGAGAGAGGGCAAAGAGGCAGCAAGAGATTTGTCTGGGGATCCA
GAAACCCATGATACCCTACTGAACACCGAATCCCCTGGAAGCCCACAGAGACAGAGACAGCAAGA
GAAGCAGAGATAAATACACTCACGCCAGGAGCTCGCTCGCTCTCTCTCTCTCTCTCACTCCTC
CCTCCCTCTCTCTCTGCCTGTCTTAGTCCTCTAGTCCTCAAATCCCAGTCCCCTGCACCCCTTC
CTGGGACACTATGTTGTCTCCGCCCTCCTGTGGAGGTGATTGGATCCTGGCTGCAGATGGGG
GTCAACACTGGACGTATGAGGGCCACATGGTCAGGACCATTGGCCAGCCTCTTACCCTGAGTGT
GGAAACAATGCCCAGTCGCCCATCGATATTCAGACAGACAGTGTGACATTTGACCCTGATTTGCC
TGCTCTGCAGCCCCACGGATATGACCAGCCTGGCACCGAGCCTTTGGACCTGCACAACAATGGCC
ACACAGTGAACCTCTCTCTGCCCTCTACCCTGTATCTGGGTGGACTTCCCCGAAAATATGTAGCT
GCCCAGCTCCACCTGCACTGGGGTCAGAAAGGATCCCCAGGGGGGTGAGAACACCAGATCAACAG
TGAAGCCACATTTGCAGAGCTCCACATTGTACATTATGACTCTGATTCTATGACAGCTTGAGTG
AGGCTGCTGAGAGGCCTCAGGGCCTGGCTGTCTGGGCATCCTAATTGAGGTGGGTGAGACTAAG
AATATAGCTTATGAACACATTCTGAGTCACCTGCATGAAGTCAGGCATAAAGATCAGAAGACCTC
AGTGCCCTCCCTTCAACCTAAGAGAGCTGCTCCCCAAACAGCTGGGGCAGTACTTCCGCTACAATG
GCTCGCTCACAACCCCCCTTGCTACCAGAGTGTGCTCTGGACAGTTTTTTATAGAAGGTCCCAG
ATTTCAATGGAACAGCTGGAAAAGCTTCAGGGGACATTGTTCTCCACAGAAGAGAGCCCTCTAA
GCTTCTGGTACAGAACTACCGAGCCCTTCAGCCTCTCAATCAGCGCATGGTCTTTGCTTCTTTCA
TCCAAGCAGGATCCTCGTATACCACAGGTGAAATGCTGAGTCTAGGTGTAGGAATCTTGGTTGGC
TGTCTCTGCCTTCTCCTGGCTGTTTATTTTCATTGCTAGAAAGATTCGGAAGAAGAGGCTGGAAAA
CCGAAAGAGTGTGGTCTTCACCTCAGCACAAGCCACGACTGAGGCATTAAATTCCTTCTCAGATAC
CATGGATGTGGATGACTTCCCTTCATGCCTATCAGGAAGCCTCTAAAATGGGGTGTAGGATCTGG
CCAGAAACACTGTAGGAGTAGTAAGCAGATGTCTCCTTCCCCTGGACATCTCTTAGAGAGGAAT
GGACCCAGGCTGTCAATCCAGGAAGAACTGCAGAGCCTTCAGCCTCTCCAAACATGTAGGAGGAA
ATGAGGAAATCGCTGTGTTGTTAATGCAGAGANCAAACTCTGTTTAGTTGCAGGGGAAGTTTGGG
ATATACCCCAAAGTCTCTACCCCTCACTTTTATGGCCCTTCCCTAGATATACTGCGGGATCT
CTCCTTAGGATAAAGAGTTGCTGTTGAAGTTGTATATTTTGTATCAATATATTTGGAAATTAAAG
TTTCTGACTTT

FIGURE 74

MLFSALLLEVIWILAADGGQHWTYEGPHGQDHWPAASYPECGNNAQSPIDIQTDSVTFDPLPALQ
PHGYDQPGTEPLDLHNNGHTVQLSLPSTLYLGGLPRKYVAAQLHLHWGQKGSPPGGSEHQINSEAT
FAELHIVHYDSDSYDSLSEAAERPQGLAVLGILIEVGETKNIAYEHLHLHEVRHKDQKTSVPP
FNLRELLPKQLGQYFRYNGSLTTPCYQSVLWTVFYRRSQISMEQLEKLQGTLFSTEEEPSKLLV
QNYRALQPLNQRMVFASFIQAGSSYTTGEMLSLGVGILVGCLCLLLAVYFIARKIRKKRLNRKS
VVFTSAQATTEA

Important features of the protein:

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 291-310

N-glycosylation site.

amino acids 213-216

Eukaryotic-type carbonic anhydrases proteins

amino acids 197-245, 104-140, 22-69

FIGURE 75

TGCCGCTGCCGCGCTGCTGCTGTTGCTCCTGGCGGCGCCTTGGGGACGGGCAGTTCCCTGTGTC
TCTGGTGGTTTGCCTAAACCTGCAAACATCACCTTCTTATCCATCAACATGAAGAATGTCCTACA
ATGGACTCCACCAGAGGGTCTTCAAGGAGTTAAAGTTACTTACACTGTGCAGTATTTTCATCACAA
ATTGGCCCACCAGAGGTGGCACTGACTACAGATGAGAAGTCCATTTCTGTTGTCCTGACAGCTCC
AGAGAAGTGGGAAGAGAAATCCAGAAGACCTTCCTGTTTCCATGCAACAAATATACTCCAATCTGA
AGTATAACGTGTCTGTGTTGAATACTAAATCAAACAGAACGTGGTCCCAGTGTGTGACCAACCAC
ACGCTGGTGTCTCACCTGGCTGGAGCCGAACACTCTTTACTGCGTACACGTGGAGTCCCTTCGTCCC
AGGGCCCCCTCGCCGTGCTCAGCCTTCTGAGAAGCAGTGTGCCAGGACTTTGAAAGATCAATCAT
CAGAGTTCAAGGCTAAAATCATCTTCTGGTATGTTTTGCCCATATCTATTACCGTGTTCCTTTTT
TCTGTGATGGGCTATTCCATCTACCGATATATCCACGTTGGCAAAGAGAAAACCCAGCAAATTT
GATTTTGATTTATGGAATGAATTTGACAAAAGATTCTTTGTGCCTGCTGAAAAATCGTGATTA
ACTTTATCACCTCAATATCTCGGATGATTCTAAAATTTCTCATCAGGATATGAGTTTACTGGGA
AAAAGCAGTGATGTATCCAGCCTTAATGATCCTCAGCCCAGCGGGAACCTGAGGCCCCCTCAGGA
GGAAGAGGAGGTGAAACATTTAGGGTATGCTTCGCATTTGATGGAAATTTTTTGTGACTCTGAAG
AAAACACGGAAGTACTTCTCTCACCCAGCAAGAGTCCCTCAGCAGAACAATACCCCCGGATAAA
ACAGTCATTGAATATGAATATGATGTGAGAACCACTGACATTTGTGCGGGCCCTGAAGAGCAGGA
GCTCAGTTTGCAGGAGGAGGTGTCCACACAAGGAACATTATTGGAGTCGCAGGCAGCGTTGGCAG
TCTTGGGCCCCGAAACGTTACAGTACTCATACACCCCTCAGCTCCAAGACTTAGACCCCTGGCG
CAGGAGCACACAGACTCGGAGGAGGGGCCGGAGGAAGAGCCATCGACGACCCTGGTCGACTGGGA
TCCCCAAACTGGCAGGCTGTGTATTCTTCGCTGTCCAGCTTCGACCAGGATTCAGAGGGCTGCG
AGCCTTCTGAGGGGATGGGCTCGGAGAGGAGGGTCTTCTATCTAGACTCTATGAGGAGCCGGCT
CCAGACAGGCCACCAGGAGAAAATGAAACCTATCTCATGCAATTCATGGAGGAATGGGGTTATA
TGTGCAGATGGAAAACTGATGCCAACACTTCCTTTTGCCTTTTGTTCCTGTGCAACAAGTGAG
TCACCCCTTTGATCCCAGCCATAAAGTACCTGGGATGAAAGAAGTTTTTCCAGTTTGTGAGTGT
CTGTGAGAATTACTTATTTCTTTCTCTATTCTCATAGCACGTGTGTGATTGGTTCATGCATGTA
GGTCTCTTAACAATGATGGTGGGCTCTGGAGTCCAGGGGCTGGCCGGTTGTTCTATGCAGAGAA
AGCAGTCAATAAATGTTTGCCAGACTGGGTGCAGAATTTATTAGGTGGGTGT

FIGURE 76

MSYNGLHQRVFKEKLLTLCSISSQIGPPEVALTTDEKSISVVLTAPEKWKRNPEDLPVSMQQIY
SNLKYNVSVLNTKSNRTWSQCVTNHTLVLTWLEPNTLYCVHVESFVPGPPRAQPSEKQCARTLK
DQSSEFKAKIIFWYVLPISITVFLFSVMGYSIYRIHVGKEKHPANLILYGNEDKRFVPAEK
IVINFITLNISSDKISHQDMSLLGKSSDVSSLNDPQPSGNLRPPQEEEEVKHLGYASHLMEIFC
DSEENTEGTSLTQQESLSRTIPDPKTVIEYEYDVRTTDCAGPEEQELSLQEEVSTQGTLLSQA
ALAVLGPQTLQYSYTPQLQDLPLAQEHTDSEEGPEEEPSTTLVDWDPQTGRLCIPSLSSFDQDS
EGCEPSEGDGLGEEGLLSRLYEPPAPDRPPGENETYLMQFMEEWGLYVQMEN

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 140-163

N-glycosylation sites.

amino acids 71-74, 80-83, 89-92, 204-207, 423-426

FIGURE 77

GAGGAGCGGGCCGAGGACTCCAGCGTGCCGAGGTCTGGCATCCTGCACTTGCTGCCCTCTGACAC
CTGGGAAGATGCGCGCCCGTGGACCTTCACCTTCTCTGTGGTTTGCTGGCAGCCACCTTGATC
CAAGCCACCTCAGTCCCCTGCACTTCTCATCCTCGGCCAAAAGTCATCAAAGAAAAGCTGAC
ACAGGAGCTGAAGGACCACAACGCCACCAGCATCCTGCAGCAGCTGCCGCTGCTCAGTGCCATGC
GGGAAAAGCCAGCCGGAGGCATCCCTGTGCTGGGCAGCCTGGTGAACACCGTCCTGAAGCACATC
ATCTGGCTGAAGGTCATCACAGCTAACATCCTCCAGCTGCAGGTGAAGCCCTCGGCCAATGACCA
GGAGCTGCTAGTCAAGATCCCCCTGGACATGGTGGCTGGATTCAACACGCCCTGGTCAAGACCA
TCGTGGAGTTCACATGACGACTGAGGCCCAAGCCACCATCCGCATGGACACCAGTGAAGTGGC
CCCACCCGCTGGTCCCTCAGTGACTGTGCCACCAGCCATGGGAGCCTGCGCATCCAATGCTGTA
TAAGCTCTCCTTCTGGTGAACGCCTTAGCTAAGCAGGTGATGAACCTCCTAGTGCCATCCCTGC
CCAATCTAGTGAAAAACCAGCTGTGTCCCGTGATCGAGGCTTCCTTCAATGGCATGTATGCAGAC
CTCCTGCAGCTGGTGAAGGTGCCCATTTCCCTCAGCATTGACCGTCTGGAGTTTGACCTTCTGTA
TCCTGCCATCAAGGGTGACACCATTAGCTCTACCTGGGGCCAAGTTGTTGGACTCACAGGGAA
AGGTGACCAAGTGGTTCAATAACTCTGCAGCTTCCCTGACAATGCCCCACCTGGACAACATCCCG
TTCAGCCTCATCGTGAGTCAGGACGTGGTGAAGCTGCAGTGGCTGCTGTGCTCTCTCCAGAAGA
ATTCATGGTCCCTGTTGGACTCTGTGCTTCCCTGAGAGTGCCCATCGGCTGAAGTCAAGCATCGGGC
TGATCAATGAAAAGGCTGCAGATAAGCTGGGATCTACCCAGATCGTGAAGATCCTAACTCAGGAC
ACTCCCGAGTTTTTTATAGACCAAGGCCATGCCAAGGTGGCCCAACTGATCGTGCTGGAAGTGT
TCCCTCCAGTGAAGCCCTCCGCCCTTTGTTACCCCTGGGCATCGAAGCCAGCTCGGAAGCTCAGT
TTTACACCAAAGGTGACCAACTTATACTCAACTTGAATAACATCAGCTCTGATCGGATCCAGCTG
ATGAACTCTGGGATTGGCTGGTTCCAACCTGATGTTCTGAAAAACATCATCACTGAGATCATCCA
CTCCATCCTGCTGCCGAACCAGAATGGCAAATTAAGATCTGGGGTCCCAGTGTCATTGGTGAAGG
CCTTGGGATTCGAGGCAGCTGAGTCTCACTGACCAAGGATGCCCTTGTGCTTACTCCAGCCTCC
TTGTGGAAACCCAGCTCTCCTGTCTCCAGTGCAAAGACTTGGATGGCAGCCATCAGGAAGGCTGG
GTCCCAGCTGGGAGTATGGGTGTGAGCTCTATAGACCATCCCTCTCTGCAATCAATAAACACTTG
CCTGTGAAAAA

FIGURE 78

MAGPWTFLLCGLLAATLIQATLSPTAVLILGPKVIEKLTQELKDHNATSILQQLPLLSAMREK
PAGGIPVLGSLVNTVLKHIIWLKVITANILQLQVKPSANDQELLVKIPLDMVAGFNTPLVKTIVE
FHMTTEAQATIRMDTSASGPTRLVLSDCATSHGSLRIQLLYKLSFLVNALAKQVMNLLVPSLPNL
VKNQLCPVIEASFNGMYADLLQLVKVPISLSIDRLEFDLLYPAIKGDITQLYLGAKLLDSQGKVT
KWFNNSAASLTMPITLDNIPFSLIVSQDVVKAAVAVALSPEEFMVLLDSVLPESAHLKSSIGLIN
EKAADKLGSTQIVKILTQDTPEFFIDQGHAKVAQLIVLEVFPSSALRPLFTLGIEASSEAQFYT
KGDQLILNLNLISSDRIQLMNSGIGWFQPDVLKNIITEIIHSILLPNQNGKLRSGVPVSLVKALG
FEAAESSLTKDALVLTASLWKPSSPVSQ

Important features of the protein:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 48-51, 264-267, 401-404

Glycosaminoglycan attachment site.

amino acids 412-415

LBP / BPI / CETP family proteins.

amino acids 407-457

FIGURE 79

GAGAGAAGTCAGCCTGGCAGAGAGACTCTGAAATGAGGGATTAGAGGTGTTCAAGGAGCAAGAGC
TTCAGCCTGAAGACAAGGGAGCAGTCCCTGAAGACGCTTCTACTGAGAGGTCTGCCATGGCCTCT
CTTGGCCTCCAACCTGTGGGCTACATCCTAGGCCTTCTGGGGCTTTTGGGCACACTGGTTGCCAT
GCTGCTCCCCAGCTGGAAAACAAGTTCTTATGTCGGTGCCAGCATTGTGACAGCAGTTGGCTTCT
CCAAGGGCCTCTGGATGGAATGTGCCACACACAGCACAGGCATCACCCAGTGTGACATCTATAGC
ACCCTTCTGGGCCTGCCCGCTGACATCCAGGCTGCCAGGCCATGATGGTGACATCCAGTGCAAT
CTCCTCCCTGGCCTGCATTATCTCTGTGGTGGGCATGAGATGCACAGTCTTCTGCCAGGAATCCC
GAGCCAAAGACAGAGTGGCGGTAGCAGGTGGAGTCTTTTTCATCCTTGGAGGCCTCCTGGGATTCT
ATTCTGTGCTGGAATCTTCATGGGATCCTACGGGACTTCTACTCACCCTGGTGCCTGACAG
CATGAAATTTGAGATTGGAGAGGCTCTTTACTTGGGCATTATTTCTTCCCTGTTCTCCCTGATAG
CTGGAATCATCCTCTGCTTTTCTGCTCATCCAGAGAAATCGCTCCAACCTACTACGATGCCTAC
CAAGCCCAACCTCTTGCCACAAGGAGCTCTCCAAGGCCTGGTCAACCTCCCAAAGTCAAGAGTGA
GTTCAATTCTACAGCCTGACAGGGTATGTGTGAAAGAACCAGGGGCCAGAGCTGGGGGTGGCTG
GGTCTGTGAAAAACAGTGGACAGCACCCCGAGGGCCACAGGTGAGGGACACTACCCTGGATCGT
GTCAGAAGGTGCTGCTGAGGATAGACTGACTTTGGCCATTGGATTGAGCAAAGGCAGAAATGGGG
GCTAGTGTAACAGCATGCAGGTTGAATTGCCAAGGATGCTCGCCATGCCAGCCTTTCTGTTTTCC
TCACCTTGCTGCTCCCCTGCCCTAAGTCCCCAACCCCTCAACTTGAAACCCCATTCCTTAAGCCA
GGACTCAGAGGATCCCTTTGCCCTCTGGTTTACCTGGGACTCCATCCCCAAACCCACTAATCACA
TCCCACTGACTGACCCTCTGTGATCAAAGACCCTCTCTCTGGCTGAGGTTGGCTCTTAGCTCATT
GCTGGGGATGGGAAGGAGAAGCAGTGGCTTTTGTGGGCATTGCTCTAACCTACTTCTCAAGCTTC
CCTCCAAAGAACTGATTGGCCCTGGAACCTCCATCCCCTCTTGTATGACTCCACAGTGTCCA
GACTAATTTGTGCATGAACTGAAATAAAACCATCCTACGGTATCCAGGGAACAGAAAGCAGGATG
CAGGATGGGAGGACAGGAAGGCAGCCTGGGACATTTAAAAAATA

FIGURE 8o

MASLGLQLVGYILGLLGLLGLTLVAMLLPSWKTSSYVGASIVTAVGFSKGLWMECATHSTGITQCD
IYSTLLGLPADIQAAQAMMVTSSAISLACIISVVGMRCTVFCQESRAKDRVAVAGGVFFILGGL
LGFIPVAVNHLHGILRDFYSPLVPDSMKFEIGEALYLGIISSLFSLIAGIILCFSCSSQRNRSNYY
DAYQAQPLATRSSPRPGQPPKVKSEFNYSYSLTGYV

Important features of the protein:

Signal peptide:

amino acids 1-24

Transmembrane domains:

amino acids 82-102, 117-140, 163-182

N-glycosylation site.

amino acids 190-193

PMP-22 / EMP / MP20 family proteins.

amino acids 46-59

FIGURE 81

CCCACGCGTCCGCGCCTCTCCCTTCTGCTGGACCTTCCTTCGTCTCTCCATCTCTCCCTCCTTTT
CCCGCGTTCTCTTTCCACCTTTCTCTTCTCCACCTTAGACCTCCCTTCCTGCCCTCCTTTCTT
GCCCACCGCTGCTTCCTGGCCCTTCTCCGACCCCGCTCTAGCAGCAGACCTCCTGGGGTCTGTGG
GTTGATCTGTGGCCCTGTGCCTCCGTGTCTTTTCGTCTCCCTTCCTCCCGACTCCGCTCCCGG
ACCAGCGGCCTGACCTGGGGAAGGATGGTTCCCGAGGTGAGGGTCTCTCCTCCTTGTGGGA
CTCGCGCTGCTCTGGTTCCCCCTGGACTCCCACGCTCGAGCCCGCCAGACATGTTCTGCCTTTT
CCATGGGAAGAGATACTCCCCGCGAGAGCTGGCACCCCTACTTGGAGCCACAAGGCCTGATGT
ACTGCCTGCGCTGTACCTGCTCAGAGGGCGCCCATGTGAGTTGTTACCGCCTCCACTGTCCGCTT
GTCCACTGCCCCAGCCTGTGACGGAGCCACAGCAATGCTGTCCAAGTGTGTGGAACCTCACAC
TCCCTCTGGACTCCGGGCCCCACCAAAGTCTGCCAGCACAAACGGGACCATGTACCAACACGGAG
AGATCTTCAGTGCCCATGAGCTGTTCCTTCCCGCTGCCCAACCAGTGTGTCTCTGCAGCTGC
ACAGAGGGCCAGATCTACTGCGGCCTCACAACCTGCCCGAACCAGGTGCCAGCACCCCTCCC
ACTGCCAGACTCCTGCTGCCAAGCCTGCAAAGATGAGGCAAGTGAAGAGGACA
GTGTGCAGTCGCTCCATGGGGTGAACATCTCAGGATCCATGTTCCAGTGATGCTGGGAGAAAG
AGAGGCCCCGGGCACCCAGCCCCACTGGCCTCAGCGCCCTCTGAGCTTCATCCCTCGCCACTT
CAGACCCAAGGGAGCAGGCAGCACAACTGTCAAGATCGTCTGAAGGAGAAACATAAGAAAGCCT
GTGTGCATGGCGGAAGACGTACTCCACGGGGAGGTGTGGCACCCGGCCTTCCGTGCCTTCGGC
CCCTTGCCCTGCATCCTATGCACCTGTGAGGATGGCCGCCAGGACTGCCAGCGTGTGACCTGTCC
CACCGAGTACCCCTGCCGTACCCCGAGAAAGTGGCTGGGAAGTGTGCAAGATTGCCCAGAGG
ACAAAGCAGACCCCTGGCCACAGTGAGATCAGTTCTACCAGGTGTCCAAGGCACCGGGCCGGGTC
CTCGTCCACACATCGGTATCCCCAAGCCCAGACAACCTGCGTCGCTTTGCCCTGGAACACGAGGC
CTCGGACTTGGTGGAGATCTACCTCTGGAAGCTGGTAAAAGATGAGGAACTGAGGCTCAGAGAG
GTGAAGTACCTGGCCCAAGGCCACACAGCCAGAATCTTCCACTTGACTCAGATCAAGAAAGTCAG
GAAGCAAGACTTCCAGAAAGAGGCACAGCACTTCCGACTGCTCGCTGGCCCCACGAAGGTCACT
GGAACGTCTTCTAGCCCAGACCCTGGAGCTGAAGGTCACGGCCAGTCCAGACAAAGTGACCAAG
ACATAACAAAGACCTAACAGTTGCAGATATGAGCTGTATAATTGTTGTTATTATATATATAATAA
TAAGAAGTTGCATTACCCTCAAAAAAAAAAAAAAAAAAAAAA

FIGURE 82

MVPEVRVLSSLLGLALLWFPLDSHARARPDMFCLFHGKRYSPGESWHPYLEPQGLMYCLRCTCSE
GAHVSCYRLHCPPVHCPQPVTEPQQCCPKCEPHTPSGLRAPPKSCQHNGTMYQHGEIFSAHELF
PSRLPNQCVCSCTEGQIYCGLTTCPEPGCPAPLPLPDSCCQACKDEASEQSDEEDSVQSLHGVR
HPQDPCSSDAGRKRGPPTAPTGLSAPLSFIPRHRPKGAGSTTVKIVLKEKHKKACVHGKGTYS
HGEVWHPAFRAFGPLPCILCTCEDGRQDCQRVTCPTTEYPCRHPKVKAGKCKICPEDKADPGHSE
ISSTRCPKAPGRVLVHTSVSPSPDNLRFALEHEASDLVEIYLWKLVKDEETEAQRGEVPGPRPH
SQNLPLDSDQESQEARLPERGTALPTARWPPRRSLERLPSPDPAEGHGQSRQSDQDITKT

Signal peptide:

amino acids 1-25

FIGURE 83

GACAGCTGTGTCTCGATGGAGTAGACTCTCAGAACAGCGCAGTTTGGCCCTCCGCTCACGCAGAGCCTCTCC
 GTGGCTTCCGCACCTTGAGCATTAGGCCAGTTCTCTCTTCTCTCTAATCCATCCGTCACCTCTCCTGTCA
 TCCGTTTCCATGCCGTGAGGTCCATTACAGAACACATCCATGGCTCTCATGCTCAGTTTGGTTCTGAGTC
 TCCTCAAGCTGGGATCAGGGCAGTGGCAGGTGTTTGGGCCAGACAAGCCTGTCCAGGCCTTGGTGGGGGAG
 GACGCAGCATTTCTCTGTTTCTGTCTCCTAAGACCAATGCAGAGGCCATGGAAGTGCGGTTCTTCAGGGG
 CCAGTTCTCTAGCGTGGTCCACCTCTACAGGGACGGGAAGGACCAGCCATTTATGCAGATGCCACAGTATC
 AAGGCAGGACAAAACCTGGTGAAGGATTCTATTGCGGAGGGGCGCATCTCTCTGAGGCTGGAACATTACT
 GTGTTGGATGCTGGCCTCTATGGGTGCAGGATTAGTTCCAGTCTTACTACCAGAAGGCCATCTGGGAGCT
 ACAGGTGTCTAGCACTGGGCTCAGTTCTCTCATTTCATCACGGGATATGTTGATAGAGACATCCAGCTAC
 TCTGTCTAGTCTCTCGGGCTGGTTCCCCCGCCACAGCGAAGTGGAAGGTCCACAAGGACAGGATTGTGCC
 ACAGACTCCAGGACAAACAGAGACATGCATGGCCTGTTTGATGTGGAGATCTCTCTGACCGTCCAAGAGAA
 CGCCGGGAGCATATCTGTTCCATGCGGCATGCTCATCTGAGCCGAGAGGTGGAATCCAGGTACAGATAG
 GAGATACCTTTTTTCGAGCCTATATCGTGGCACCTGGCTACCAAAGTACTGGGAATACTCTGCTGTGGCCTA
 TTTTTTGGCATTGTTGGACTGAAGATTTTCTCTCCAAATTCAGTGGAAAATCCAGGCGGAAGTGGACTG
 GAGAAGAAAGCACGGACAGGCAGAATTGAGAGACGCCCCGAAACACGCAGTGGAGGTGACTCTGGATCCAG
 AGACGGCTCACCCGAAGCTCTGCGTTTCTGATCTGAAAACGTAAACCATAGAAAAGCTCCCCAGGAGGTG
 CCTCACTCTGAGAAGAGATTTACAAGGAAGAGTGTGGTGGCTTCTCAGAGTTTCCAAGCAGGGAAACATTA
 CTGGGAGGTGGACGGAGGACACAATAAAAGGTGGCGCTGGGAGTGTGCCGGATGATGTGGACAGGAGGA
 AGGAGTACGTGACTTTGTCTCCCGATCATGGGTACTGGGTCTCAGACTGAATGGAGAACATTGTATTTTC
 ACATTAAATCCCCGTTTATCAGCGTCTTCCCCAGGACCCACCTACAAAAATAGGGGTCTTCTGGACTA
 TGAGTGTGGGACCATCTCCTTCTTCAACATAAATGACCAGTCCCTTATTTATACCCGTGACATGTGGTTTG
 AAGGCTTATTGAGGCCCTACATTGAGTATCCGTCTATAATGAGCAAAATGGAAGTCCCATAGTCATCTGC
 CCAGTCACCCAGGAATCAGAGAAAGAGGCCTCTTGGCAAAGGGCCTCTGCAATCCCAGAGACAAGCAACAG
 TGAGTCTCTCTCACAGGCAACCACGCCCTTCTCTCCAGGGGTGAAATGTAGGATGAATCACATCCCACAT
 TCTTCTTTAGGGATATTAAAGTCTCTCTCCAGATCCAAAGTCCCGCAGCAGCCGGCCAAGGTGGCTTCCA
 GATGAAGGGGGACTGGCCTGTCCACATGGGAGTCAGGTGTCTGAGCTGCCCTGAGCTGGGAGGGAAGAAGG
 CTGACATTACATTTAGTTTGTCTCTCACTCCATCTGGCTAAGTGATCTTGAAATACCACCTCTCAGGTGAAG
 AACCCTCAGGAATCCCATCTCACAGGCTGTGGTGTAGATTAAGTAGACAAGGAATGTGAATAATGCTTAG
 ATCTTATTGATGACAGAGTGTATCCTAATGGTTTGTTCATTATATTACACTTTCAGTAAAAAAA

FIGURE 84

MALMLSLVLSELLKLGGQWQVFGPDKPVQALVGEDAAFSCFLSPKTNAEAMEVRFFRGQFSSVVH
LYRDGKDQPFMQMPQYQGRTKLVKDSIAEGRISLRLENITVLDAGLYGCRISSQSYQKAIWELQ
VSALGSVPLISITGYVDRDIQLLCQSSGWFPPTAKWKGPQGQDLSTDSRTNRDMHGLFDVEISL
TVQENAGSISCSMRHAHLSREVESRVQIGDTFFEPISWHLATKVLGILCCGLFFGIVGLKIFFSK
FQWKIQAELDWRRKHGQAEIRDARKHAVEVTLPETAHPKLCVSDLKTVTHRKAPQEVPHSEKRF
TRKSVVASQS FQAGKHYWEVDGGHNKRWRVGVCRDDVDRRKEYVTLSPDHGYWVLRNLNGEHLFT
LNPRFISVFRTPTPTKIGVFLDYECGTISFFNINDQSLIYTLTCRFEGLLRPYIEYPSYNEQNGT
PIVICPVTQESEKEASWQRASAIPESTNSSESSQATTFFLPRGEM

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 239-255

FIGURE 85

AACAGACGTTCCCTCGCGGCCCTGGCACCTCTAACCCAGACATGCTGCTGCTGCTGCTGCCCCCT
GCTCTGGGGGAGGGAGAGGGCGGAAGGACAGACAAGTAAACTGCTGACGATGCAGAGTTCCGTGA
CGGTGCAGGAAGGCCTGTGTGTCCATGTGCCCTGCTCCTTCTCCTACCCCTCGCATGGCTGGATT
TACCCCTGGCCCAGTAGTTTCATGGCTACTGGTTCGGGAAGGGGCCAATACAGACCAGGATGCTCC
AGTGGCCACAAACAACCCAGCTCGGGCAGTGTGGGAGGAGACTCGGGACCGATTCCACCTCCTTG
GGGACCCACATACCAAGAATTGCACCCTGAGCATCAGAGATGCCAGAAGAAGTGATGCGGGGAGA
TACTTCTTTCGTATGGAGAAAGGAAGTATAAAATGGAATTATAAAACATCACCGGCTCTCTGTGAA
TGTGACAGCCTTGACCCACAGGCCCAACATCCTCATCCCAGGCACCCTGGAGTCCGGCTGCCCCC
AGAATCTGACCTGCTCTGTGCCCTGGGCCTGTGAGCAGGGGACACCCCTATGATCTCCTGGATA
GGGACCTCCGTGTCCCCCTGGACCCCTCCACCACCGCTCCTCGGTGCTCACCCCTCATCCCACA
GCCCCAGGACCATGGCACCAGCCTCACCTGTCAGGTGACCTTCCCTGGGGCCAGCGTGACCACGA
ACAAGACCGTCCATCTCAACGTGTCTACCCGCTCAGAACTTGACCATGACTGTCTTCCAAGGA
GACGGCACAGTATCCACAGTCTTGGGAAATGGTTCATCTCTGTCACTCCCAGAGGGCCAGTCTCT
GCGCCTGGTCTGTGCAGTTGATGCAGTTGACAGCAATCCCCCTGCCAGGCTGAGCCTGAGCTGGA
GAGGCCTGACCCTGTGCCCTCACAGCCCTCAAACCCGGGGGTGCTGGAGCTGCCTTGGGTGCAC
CTGAGGGATGCAGCTGAATTCACCTGCAGAGCTCAGAACCCTCTCGGCTCTCAGCAGGTCTACCT
GAACGTCTCCCTGCAGAGCAAAGCCACATCAGGAGTGACTCAGGGGGTGGTCGGGGGAGCTGGAG
CCACAGCCCTGGTCTTCTGTCTTCTGCGTCATCTTCGTTGTAGTGAGGTCTGCAGGAAGAA
TCGGCAAGGCCAGCAGCGGGCGTGGGAGATACGGGCATAGAGGATGCAAACGCTGTCAGGGGTTT
AGCCTCTCAGGGGCCCTGACTGAACCTTGGGCAGAAGACAGTCCCCCAGACCAGCTCCCCCAG
CTTCTGCCCCGCTCCTCAGTGGGGGAAGGAGAGCTCCAGTATGCATCCCTCAGCTTCCAGATGGTG
AAGCCTTGGGACTCGCGGGGACAGGAGGCCACTGACACCGAGTACTCGGAGATCAAGATCCACAG
ATGAGAACTGCAGAGACTCACCTGATTGAGGGATCACAGCCCCTCCAGGCAAGGGAGAAGTCA
GAGGCTGATTCTTGTAGAATTAACAGCCCTCAACGTGATGAGCTATGATAACACTATGAATTATG
TGCAGAGTGAAAAGCACACAGGCTTTAGAGTCAAAGTATCTCAAACCTGAATCCACACTGTGCCC
TCCCTTTTATTTTTTTAACTAAAAGACAGACAAATTCCTA

FIGURE 86

MLLLLLPLLWGRERAEGQTSKLLTMQSSVTVQEGLCVHVPCSFSPSHGWIYPGPVVHGYWFREG
ANTDQDAPVATNNPARAVWEETRDRFHLLGDPHTKNCTLSIRDARRSDAGRYFFRMEKGSIKWNY
KHHRLSVNVTALTHRPNILIPGTLESGCPQNLTCVWPWACEQGTPPMISWIGTSVSPLDPSTTRS
SVLTLIPQPQDHGTS LTCQVT FPGASVT TTKTVHLNVSYPQNLTMTVFQGDGT VSTVLGNGSSL
SLPEGQSLRLVCAVDAVDSNPPARLSLSWRGLTLCPSQPSNPGVLELPWVHLRDAAEFTCRAQNP
LGSQQVYLVNLSLQSKATSGVTQGVVGGAGATALVFLSFCVIFVVVRSCRKKSARPAAGVGD TGIE
DANAVRG SASQGPLEPWAEDSPPDQPPPASARSSVGE GELQYASLSFQMKPWDSRGQEATDTE
YSEIKIHR

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 351-370

FIGURE 87

AGAAAGCTGCACTCTGTTGAGCTCCAGGGCGCAGTGGAGGGAGGGAGTGAAGGAGCTCTCTGTAC
CCAAGGAAAGTGCAGCTGAGACTCAGACAAGATTACAATGAACCAACTCAGCTTCCTGCTGTTTC
TCATAGCGACCACCAGAGGATGGAGTACAGATGAGGCTAATACTTACTTCAAGGAATGGACCTGT
TCTTCGTCTCCATCTCTGCCCAGAAGCTGCAAGGAAATCAAAGACGAATGTCCTAGTGCATTTGA
TGGCCTGTATTTTCTCCGCACTGAGAATGGTGTATCTACCAGACCTTCTGTGACATGACCTCTG
GGGGTGGCGGCTGGACCCTGGTGGCCAGCGTGCATGAGAATGACATGCGTGGGAAGTGCACGGTG
GGCGATCGCTGGTCCAGTCAGCAGGGCAGCAAAGCAGACTACCCAGAGGGGACGGCAACTGGGC
CAACTACAACACCTTTGGATCTGCAGAGGCGGCCACGAGCGATGACTACAAGAACCCTGGCTACT
ACGACATCCAGGCCAAGGACCTGGGCATCTGGCACGTGCCCAATAAGTCCCCCATGCAGCACTGG
AGAAACAGCTCCCTGCTGAGGTACCGCACGGACACTGGCTTCCTCCAGACACTGGGACATAATCT
GTTTGGCATCTACCAGAAATATCCAGTGAAATATGGAGAAGGAAAGTGTGGACTGACAACGGCC
CGGTGATCCCTGTGGTCTATGATTTTGGCGACGCCAGAAAACAGCATCTTATTACTCACCCAT
GGCCAGCGGAATTCAGTGCGGGATTGTTCAGTTCAGGGTATTTAATAACGAGAGAGCAGCCAA
CGCCTTGCTGTGCTGGAATGAGGGTCACCGGATGTAACACTGAGCATCACTGCATTGGTGGAGGAG
GATACTTTCCAGAGGCCAGTCCCCAGCAGTGTGGAGATTTTCTGGTTTTGATTGGAGTGGATAT
GGAACATCATGTTGGTTACAGCAGCAGCCGTGAGATAACTGAGGCAGCTGTGCTTCTATTCTATCG
TTGAGAGTTTTGTGGGAGGGAACCCAGACCTCTCCTCCCAACCATGAGATCCCAAGGATGGAGAA
CAACTTACCCAGTAGCTAGAATGTTAATGGCAGAAGAGAAAACAATAAATCATATTGACTCAAGA
AAAAAA

FIGURE 88

MNQLSFLLFLIATTRGWSTDEANTYFKEWTCSSSPSLPRSCKEIKDECPSAFDGLYFLRTENGVI
YQTFCDMTSGGGGWTLVASVHENDMRGKCTVGDRWSSQQGSKADYPEGDGNWANYNTFGSAEAAT
SDDYKNPGYYDIQAKDLGIWHVPNKSEPMQHWRNSSLLRYRTDTGFLQTLGHNLFGIYQKYPVKYG
EGKCWTDNGPVIPIVVYDFGDAQKTASYSPYQOREFTAGFVQFRVFNNERAANALCAGMRVTGCN
TEHHCIGGGGYFPEASPQQCGDFSGFDWSGYGTHVGYSRSSREITEAAVLLFYR

Important features:

Signal peptide:

amino acids 1-16

N-glycosylation site.

amino acids 163-167

Glycosaminoglycan attachment sites.

amino acids 74-78, 289-293

N-myristoylation sites.

amino acids 76-82, 115-121, 124-130, 253-259, 292-298

FIGURE 89

CTAGATTTGTCGGCTTGCGGGGAGACTTCAGGAGTCGCTGTCTCTGAACTTCCAGCCTCAGAGAC
CGCCGCCCTTGTCCCCGAGGGCCATGGGCCGGGTCTCAGGGCTTGTGCCCTCTCGCTTCCTGACG
CTCCTGGCGCATCTGGTGGTCGTACACCTTATTCTGGTCCCGGGACAGCAACATACAGGCCTG
CCTGCCCTCTCACGTTACCCCCGAGGAGTATGACAAGCAGGACATTAGCTGGTGGCCGCGCTCT
CTGTCACCCTGGGCCCTCTTTGCAGTGGAGCTGGCCGGTTTCCTCTCAGGAGTCTCCATGTTCAAC
AGCACCAGAGCCTCATCTCCATTGGGGCTCACTGTAGTGCATCCGTGGCCCTGTCTTCTTCAT
ATTTCGAGCGTTGGGAGTGCACCTACGTATTGGTACATTTTGTCTTCTGCAGTGCCCTTCAGCTG
TCACTGAAATGGCTTTATTTCGTCACCGTCTTTGGGCTGAAAAAGAAACCTTCTTCATTACCTTCA
TGACGGGAACCTAAGGACGAAGCCTACAGGGGCAAGGGCCGCTTCGTATTCTGGAAGAAGGAAG
GCATAGGCTTCGGTTTTCCCTCGGAACTGCTTCTGCTGGAGGATATGTGTTGGAATAATTACG
TCTTGAGTCTGGGATTATCCGCATTGTATTTAGTGCTTTGTAATAAAATATGTTTTGTAGTAACA
TTAAGACTTATATACAGTTTTAGGGGACAATTAAAAAAAAAAAA

Year	1900	1901	1902	1903	1904	1905	1906	1907	1908	1909	1910	1911	1912	1913	1914	1915	1916	1917	1918	1919	1920	1921	1922	1923	1924	1925	1926	1927	1928	1929	1930	1931	1932	1933	1934	1935	1936	1937	1938	1939	1940	1941	1942	1943	1944	1945	1946	1947	1948	1949	1950	1951	1952	1953	1954	1955	1956	1957	1958	1959	1960	1961	1962	1963	1964	1965	1966	1967	1968	1969	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100
1900	1901	1902	1903	1904	1905	1906	1907	1908	1909	1910	1911	1912	1913	1914	1915	1916	1917	1918	1919	1920	1921	1922	1923	1924	1925	1926	1927	1928	1929	1930	1931	1932	1933	1934	1935	1936	1937	1938	1939	1940	1941	1942	1943	1944	1945	1946	1947	1948	1949	1950	1951	1952	1953	1954	1955	1956	1957	1958	1959	1960	1961	1962	1963	1964	1965	1966	1967	1968	1969	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100	

Transmembrane domain:
amino acids 12-28 (type II), 51-66, 107-124

FIGURE 91

CTGGGACCCCGAAAAGAGAAGGGGAGAGCGAGGGGACGAGAGCGGAGGAGGAAGATGCAACTGAC
TCGCTGCTGCTTCGTGTTCTGCTGTCAGGGTAGCCTCTATCTGGTCATCTGTGGCCAGGATGATG
GTCCTCCCGGCTCAGAGGACCCTGAGCGTGATGACCACGAGGGCCAGCCCCGGCCCCGGGTGCCT
CGGAAGCGGGGCCACATCTCACCTAAGTCCCGCCCATGGCCAATTCCACTCTCCTAGGGCTGCT
GGCCCCGCTGGGGAGGCTTGGGGCATTCTTGGGCAGCCCCCAACCGCCCGAACCACAGCCCCC
CACCCTCAGCCAAGGTGAAGAAAATCTTTGGCTGGGGCGACTTCTACTCCAACATCAAGACGGTG
GCCCTGAACCTGCTCGTCACAGGGAAGATTGTGGACCATGGCAATGGGACCTTCAGCGTCCACTT
CCAACACAATGCCACAGGCCAGGGAAACATCTCCATCAGCCTCGTGCCCCCAGTAAAGCTGTAG
AGTTCCACCAGGAACAGCAGATCTTCATCGAAGCCAAGGCCTCCAAAATCTTCAACTGCCGGATG
GAGTGGGAGAAGGTAGAACGGGGCCCGGACCTCGCTTTGCACCCACGACCCAGCCAAGATCTG
CTCCCAGAGACCACGCTCAGAGCTCAGCCACCTGGAGCTGCTCCCAGCCCTTCAAAGTCGTCTGTG
TCTACATCGCCTTCTACAGCACGGACTATCGGCTGGTCCAGAAGGTGTGCCCAGATTACAACCTAC
CATAGTGATACCCCCTACTACCCATCTGGGTGAACCCGGGGCAGGCCACAGAGGCCAGGCCAGGGC
TGGAAGGACAGGCCTGCCCATGCAGGAGACCATCTGGACACCGGGCAGGGAAGGGGTGGGCCTC
AGGCAGGGAGGGGGGTGGAGACGAGGAGATGCCAAGTGGGGCCAGGGCCAAGTCTCAAGTGGCAG
AGAAAGGGTCCCAAGTGCTGGTCCCAACCTGAAGCTGTGGAGTGAAGTACAGAGGAGCACTGG
AGGAGGAGTGGGCTCTCTGTGCAGCCTCACAGGGCTTTGCCACGGAGCCACAGAGAGATGCTGGG
TCCCCGAGGCCTGTGGGCAGGCCGATCAGTGTGGCCCCAGATCAAGTCATGGGAGGAAGCTAAGC
CCTTGCTTCTTGCCATCCTGAGGAAAGATAGCAACAGGGAGGGGGAGATTTTCATCAGTGTGGACA
GCCTGTCAACTTAGGATGGATGGCTGAGAGGGCTTCTAGGAGCCAGTCAGCAGGGTGGGGTGGG
GCCAGAGGAGCTCTCCAGCCCTGCCTAGTGGGCGCCCTGAGCCCTTGTCTGTGCTGAGCATGG
CATGAGGCTGAAGTGGCAACCCTGGGGTCTTTGATGTCTTGACAGATTGACCATCTGTCTCCAGC
CAGGCCACCCCTTTCCAAAATTCCTCTTTGCCAGTACTCCCCCTGTACCACCCATTGCTGATG
GCACACCCATCCTTAAGCTAAGACAGGACGATTGTGGTCTCCACACTAAGGCCACAGCCCATC
CGCGTGTGTGTGTCCCTCTTCCACCCCAACCCCTGCTGGCTCCTCTGGGAGCATCCATGTCCCG
GAGAGGGGTCCCTCAACAGTCAGCCTCACCTGTGACACGGGGTCTCCCGGATCTGGATGGCGC
CGCCCTCTCAGCAGCGGGCACGGGTGGGGCGGGCGGGCCGAGAGCATGTGCTGGATCTGTTC
TGTGTGTCTGTCTGTGGGTGGGGGAGGGGAGGGAAGTCTTGTGAAACCGCTGATTGCTGACTTT
TGTGTGAAGAATCGTGTCTTGGAGCAGGAAATAAAGCTTGCCCCGGGGCA

FIGURE 92

MQLTRCCFVFLVQGSILYLVICGQDDGPPGSEDPERDDHEGQPRPRVPRKRGHISPKSRPMANSTL
LGLLAPPGEAWGILGQPPNRPNHSPPPSAKVKKIFGWGDFYSNIKTVALNLLVTGKIVDHGNGTF
SVHFQHNATGQGNISISLVPPSKAVEFHQEQQIFIEAKASKIFNCRMWEKVERGRRTSLCTHDP
AKICSRDHAQSSATWSCSQPFKVVVCVYIAFYSTDYRLVQKVC PDYNYHSDTPYYPSG

Important features of the protein:

Signal peptide:

amino acids 1-14

N-glycosylation sites.

amino acids 62-65, 127-130, 137-140, 143-146

2-oxo acid dehydrogenases acyltransferase

amino acids 61-71

FIGURE 93

CGGTGGCCATGACTGCGGCCGTGTTCTTCGGCTGCGCCTTCATTGCCTTCGGGCCTGCGCTCGCC
CTTTATGTCTTCACCATCGCCATCGAGCCGTTGCGTATCATCTTCCTCATCGCCGAGCTTTCTT
CTGGTTGGTGTCTCTACTGATTTTCGTCCCTTGTTTGGTTCATGGCAAGAGTCATTATTGACAACA
AAGATGGACCAACACAGAAATATCTGCTGATCTTTGGAGCGTTTGTCTCTGTCTATATCCAAGAA
ATGTTCCGATTTGCATATTATAAACTCTTAAAAAAGCCAGTGAAGGTTTGAAGAGTATAAACC
AGGTGAGACAGCACCCCTCTATGCGACTGCTGGCCTATGTTTCTGGCTTGGGCTTTGGAATCATGA
GTGGAGTATTTTCCTTTGTGAATACCCTATCTGACTCCTTGGGGCCAGGCACAGTGGGCATTTCAT
GGAGATTCTCCTCAATTCTTCCTTTATTTCAGCTTTCATGACGCTGGTCATTATCTTGCTGCATGT
ATTCTGGGGCATTGTATTTTTTGGATGGCTGTGAGAAGAAAAGTGGGGCATCCTCCTTATCGTTC
TCCTGACCCACCTGCTGGTGTGAGCCAGACCTTCATAAGTTCTTATTATGGAATAAACCTGGCG
TCAGCATTATAATCCTGGTGCTCATGGGCACCTGGGCATTCTTAGCTGCGGGAGGCAGCTGCCG
AAGCCTGAAACTCTGCCTGCTCTGCCAAGACAAGAAGTTTCTTCTTTACAACCAGCGCTCCAGAT
AACCTCAGGGAACCAGCACTTCCCAAACCGCAGACTACATCTTTAGAGGAAGCACAACGTGCCT
TTTTCTGAAAATCCCTTTTCTGGTGAATTGAGAAAGAAATAAACTATGCAGATA

FIGURE 94

MTAAVFFGCAFIAFGPALALYVFTIAIEPLRIIFLIAGAFFWLVSLLISSLVWFMARVIIDNKDG
PTQKYLLIFGAFVSVYIQEMFRFAYYKLLKKASEGLKSINPGETAPSMRLLAYVSGLGFGIMSGV
FSFVNTLSDSLPGTVGIHGDSPOFFLYSAFMTLVIILLHVFWGIVFFDGCEKKKWGILLIVLLT
HLLVSAQTFISSYYGINLASAFIILVLMGTWAFLAAGGSCRSLKCLLCQDKNFLLYNQRSR

Important features of the protein:

Signal peptide:

amino acids 1-19

Transmembrane domains:

amino acids 32-51, 119-138, 152-169, 216-235

Glycosaminoglycan attachment site.

amino acids 120-123

Sodium:neurotransmitter symporter family protein

amino acids 31-65

FIGURE 95

AATTTTTCACCAGAGTAAACTTGAGAAACCAACTGGACCTTGAGTATTGTACATTTTGCCTCGTG
GACCCAAAGGTAGCAATCTGAAACATGAGGAGTACGATTCTACTGTTTTGTCTTCTAGGATCAAC
TCGGTCATTACCACAGCTCAAACCTGCTTTGGGACTCCCTCCCACAAAACCTGGCTCCGGATCAGG
GAACACTACCAAACCAACAGCAGTCAAATCAGGTCTTTCCTTCTTTAAGTCTGATACCATTAACA
CAGATGCTCACACTGGGGCCAGATCTGCATCTGTTAAATCCTGCTGCAGGAATGACACCTGGTAC
CCAGACCCACCCATTGACCCTGGGAGGGTTGAATGTACAACAGCAACTGCACCCACATGTGTTAC
CAATTTTGTGCACACAACTTGGAGCCCAGGGCACTATCCTAAGCTCAGAGGAATTGCCACAAATC
TTCACGAGCCTCATCATCCATTCCCTTGTTCCTGGGAGGCATCCTGCCACCAGTCAGGCAGGGGC
TAATCCAGATGTCCAGGATGGAAGCCTTCCAGCAGGAGGAGCAGGTGTAAATCCTGCCACCCAGG
GAACCCAGCAGGCCCGCTCCCAACTCCCAGTGGCACAGATGACGACTTTGCAGTGACCACCCCT
GCAGGCATCCAAAGGAGCACACATGCCATCGAGGAAGCCACCACAGAATCAGCAAATGGAATTCA
GTAAGCTGTTTCAAATTTTTTCAACTAAGCTGCCTCGAATTTGGTGATACATGTGAATCTTTATC
ATTGATTATATTATGGAATAGATTGAGACACATTGGATAGTCTTAGAAGAAATTAATTCTTAATT
TACCTGAAAAATATTCTTGAAATTTTCAAGAAATATGTTCTATGTAGAGAATCCCAACTTTTAAAA
CAATAATTCAATGGATAAATCTGTCTTTGAAATATAACATTATGCTGCCTGGATGATATGCATAT
TAAACATATTTGGAAAAC TGAAA
AAA

FIGURE 96

MRSTILLFCLLGSTRSLPQLKPALGLPPTKLAPDQGTLPNQQQSNQVFPSLSLIPLTQM
LTLGPDHLHLLNPAAGMTPGTQTHPLTLGGLNVQQQLHPHVLPIFVTQLGAQGTLISSEE
LPQIFTSLLIHSFLFPGGILPTSQAGANPDVQDGSLPAGGAGVNPATQGTPAGRLPTPSG
TDDDFAVTTPAGIQRSTHAIEEATTESANGIQ

Signal peptide:

amino acids 1-16

FIGURE 97

GCTCAAGTGCCCTGCCCTTGCCCCACCCAGCCAGCCTGGCCAGAGCCCCCTGGAGAAGGAGCTCT
CTTCTTGCTTGGCAGCTGGACCAAGGGAGCCAGTCTTGGGCGCTGGAGGGCCTGTCTTGACCATG
GTCCCTGCCTGGCTGTGGCTGCTTTGTGTCTCCGTCCCCCAGGCTCTCCCCAAGGCCAGCCTGC
AGAGCTGTCTGTGGAAGTTCAGAAAACTATGGTGGAAATTTCCCTTTATACCTGACCAAGTTGC
CGCTGCCCCGTGAGGGGGCTGAAGGCCAGATCGTGCTGTGAGGGGACTCAGGCAAGGCAACTGAG
GGCCCATTTGCTATGGATCCAGATTCTGGCTTCCGTGCTGGTGACCAGGGCCCTGGACCGAGAGGA
GCAGGCAGAGTACCAGTACAGGTCAACCTGGAGATGCAGGATGGACATGTCTGTGGGGTCCAC
AGCCTGTGCTTGTGCACGTGAAGGATGAGAATGACCAGGTGCCCCATTTCTCTCAAGCCATCTAC
AGAGCTCGGCTGAGCCGGGGTACCAGGCCTGGCATCCCCCTTCCCTTCTCCTTGAGGCTTCAGACCG
GGATGAGCCAGGCACAGCCAACCTCGGATCTTCGATTCACATCCTGAGCCAGGCTCCAGCCAGC
CTTCCCCAGACATGTTCCAGCTGGAGCCTCGGCTGGGGGCTCTGGCCCTCAGCCCCAAGGGGAGC
ACAGCCTTGACCAACCCCTGGAGAGGACCTACCAGCTGTTGGTACAGGTCAAGGACATGGGTGA
CCAGGCCCTCAGGCCACAGGCCACTGCCACCGTGAAGTCTCCATCATAGAGAGCAGCTGGGTGT
CCCTAGAGCCTATCCACCTGGCAGAGAATCTCAAAGTCTTATACCCGCACCACATGGCCAGGTA
CACTGGAGTGGGGGTGATGTGCATATCACCTGGAGAGCCATCCCCGGGACCCTTTGAAGTGAA
TGCAGAGGGAAACCTCTACGTGACCAGAGAGCTGGACAGAGAAGCCAGGCTGAGTACCTGCTCC
AGGTGCGGGCTCAGAATTTCCATGGCGAGGACTATGCGGGCCCTCTGGAGCTGCACGTGCTGGTG
ATGGATGAGAATGACAACGTGCCATCTGCCCTCCCCGTGACCCACAGTCAGCATCCCTGAGCT
CAGTCCACCAGGTACTGAAGTGAAGTACTGACTGTGACAGAGGATGCAGATGCCCCGGCTCCCCCA
ATTCCACGTTGTGTATCAGCTCCTGAGCCCTGAGCCTGAGGATGGGGTAGAGGGGAGAGCCTTC
CAGGTGGACCCCACTTCAGGCAGTGTGACGCTGGGGGTGCTCCCACTCCGAGCAGGCCAGAACAT
CCTGCTTCTGGTGCTGGCCATGGACCTGGCAGGCGCAGAGGGTGGCTTCAGCAGCACGTGTGAAG
TCGAAGTCGAGTCACAGATATCAATGATCACGCCCTGAGTTCATCACTCCAGATTGGGCCCT
ATAAGCCTCCCTGAGGATGTGGAGCCCGGACTCTGGTGGCCATGCTAACAGCCATTGATGCTGA
CCTCGAGCCCGCTTCCGCCTCATGGATTTTGCCATTGAGAGGGGAGACACAGAAGGGACTTTTG
GCCTGGATTGGGAGCCAGACTCTGGGCATGTTAGACTCAGACTCTGCAAGAACCTCAGTTATGAG
GCAGCTCCAAGTCATGAGGTGGTGGTGGTGGTGCAGAGTGTGGCGAAGCTGGTGGGGCCAGGCCC
AGGCCCTGGAGCCACCGCCACGGTGAAGTGTGCTAGTGGAGAGAGTGTGCCACCCCCAAGTTGG
ACCAGGAGAGCTACGAGGCCAGTGTCCCATCAGTGCCCCAGCCGGCTCTTTCCTGCTGACCATC
CAGCCCTCCGACCCCATCAGCCGAACCTCAGGTTCTCCCTAGTCAATGACTCAGAGGGCTGGCT
CTGCATTGAGAAATTTCCGGGGAGGTGCACACCGCCAGTCCCTGCAGGGCGCCAGCCTGGGG
ACACCTACACGGTGCTTGTGGAGGCCAGGATACAGCCCTGACTCTTGCCCTGTGCCCTCCCAA
TACCTCTGCACACCCCGCAAGACCATGGCTTGATCGTGAGTGGACCCAGCAAGGACCCCGATCT
GGCCAGTGGGCACGGTCCCTACAGCTTACCCTTGGTCCCAACCCACGGTGAACGGGATTGGC
GCCTCCAGACTCTCAATGGTTCCCATGCCTACCTCACCTTGGCCCTGCATTGGGTGGAGCCAGT
GAACACATAATCCCCGTGGTGGTGCAGCCACAATGCCAGATGTGGCAGCTCCTGGTTTCGAGTGAT
CGTGTGTGCTGCAACGTGGAGGGGAGTGCATGCGCAAGGTGGGCCGATGAAGGGCATGCCCA
CGAAGCTGTGGCAGTGGGCATCCTTGTAGGCACCCTGGTAGCAATAGGAATCTTCTCATCCTC
ATTTTCACCCACTGGACCATGTCAAGGAAGAAGGACCCGGATCAACCAGCAGACAGCGTGCCCT
GAAGGCGACTGTCTGAATGGCCAGGCAGCTCTAGCTGGGAGCTTGGCCTCTGGCTCCATCTGAG
TCCCTGGGAGAGAGCCAGCACCCAGATCCAGCAGGGGACAGGACAGAGTAGAAGCCCTCCA
TCTGCCCTGGGGTGGAGGCACCATCACCATCACCAGGCATGTCTGCAGAGCCTGGACCAACTT
TATGGAAGTGGCCATGGGAGTGCTCCAAATGTGAGGGTGTTCGCCAATAATAAAGCCCGAGAA
CTGGGCTGGGCCCTATGGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAG

FIGURE 98

MVPAWLWLLCVSVPQALPKAQAELSVPEVNYGGNFPLYLTKLPLPREGAEGQIVLSGDSGKAT
EGPFAMDPDSGFLLVTRALDREEQAEYQLQVTLEMQDGHVLWGPQPVLVHVKDENDQVPFHSQAI
YRARLSRGTRPGIPFLFLEASDRDEPGTANSDLRFHILSQAPAQSPDMFQLEPRLGALALSPKG
STSLDHALERTYQLLVQVKMDGDQASGHQATATVEVSI IESTWVSLEPIHLAENLKVLYPHHMAQ
VHWSGGDVHYHLESHPPGPFVNAEGNLYVTRELDREAQAEYLLQVRAQNSHGEDIAAPLELHVL
VMDENDNVPICPPRDPTVSIPELSPPGTEVTRLSAEDADAPGSPNSHVYQLLSPEPEDGVEGRA
FQVDPTSGSVTLGVLPRLRAGQNILLVLAMDLAGAEGGFSSTCEVEVAVTDINDHAPEFITSQIG
PISLPEDVEPGTLVAMLTADLEPAFRLMDFAIERGDTEGTFGLDWEPDSGHVRLRLCKNLSY
EAAPSHEVVVVVQSVAKLVGPGPGGATATVTVLVERVMPPPKLDQESYEASVPI SAPAGSFLLT
IQPSDPI SRTLRFSLVNDSEGWLCIEKFSGEVHTAQSLQGAQPGDTYTVLVEAQDTALT LAPVPS
QYLCTPRQDHGLIVSGPSKDPDLASGHGPYSFTLGPNPTVQRDWRLQTLNGSHAYLT LALHWVEP
REHIIPVVVSHNAQMWQLLVRVIVCRCNVEGQCMRKVGRMKGMPTKLSAVGILVGT LVAIGIFLI
LIFTHWTMSRKKDPDQPADSVPLKATV

Signal peptide:

amino acids 1-18

Transmembrane domain:

amino acids 762-784

FIGURE 99

GGCTGACCGTGCTACATTGCCTGGAGGAAGCCTAAGGAACCCAGGCATCCAGCTGCCCACGCCTG
 AGTCCAAGATTCTTCCAGGAACACAAACGTAGGAGACCCACGCTCCTGGAAGCACCAGCCTTTA
 TCTCTTCACTTCAAGTCCCCCTTTCTCAAGAATCCTCTGTTCTTTGCCCTCTAAAGTCTTGGTAC
 ATCTAGGACCCAGGCATCTTGCTTTCAGCCACAAAGAGACAGATGAAGATGCAGAAAGGAAATG
 TTCTCCTTATGTTTGGTCTACTATTGCATTTAGAAGCTGCAACAAATTCCAATGAGACTAGCACC
 TCTGCCAACACTGGATCCAGTGTGATCTCCAGTGGAGCCAGCACAGCCACCAACTCTGGGTCCAG
 TGTGACCTCCAGTGGGGTCAGCACAGCCACCATCTCAGGGTCCAGCGTGACCTCCAATGGGGTCA
 GCATAGTCACCAACTCTGAGTTCCATAACAACCTCCAGTGGGATCAGCACAGCCACCAACTCTGAG
 TTCAGCACAGCGTCCAGTGGGATCAGCATAGCCACCAACTCTGAGTCCAGCACAACTCCAGTGG
 GGCCAGCACAGCCACCAACTCTGAGTCCAGCACACCTCCAGTGGGGCCAGCACAGTCACCAACT
 CTGGGTCCAGTGTGACCTCCAGTGGAGCCAGCACTGCCACCAACTCTGAGTCCAGCACAGTGTCC
 AGTAGGGCCAGCACTGCCACCAACTCTGAGTCTAGCACACTCTCCAGTGGGGCCAGCACAGCCAC
 CAACTCTGACTCCAGCACAACTCCAGTGGGGCTAGCACAGCCACCAACTCTGAGTCCAGCACAA
 CCTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACAGTGTCCAGTAGGGCCAGCACT
 GCCACCAACTCTGAGTCCAGCACAACTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCCAG
 AACGACCTCCAATGGGGCTGGCACAGCCACCAACTCTGAGTCCAGCACAGCTCCAGTGGGGCCA
 GCACAGCCACCAACTCTGACTCCAGCACAGTGTCCAGTGGGGCCAGCACTGCCACCAACTCTGAG
 TCCAGCACGACCTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACGACCTCCAGTGG
 GGCTAGCACAGCCACCAACTCTGACTCCAGCACAACTCCAGTGGGGCCGGCACAGCCACCAACT
 CTGAGTCCAGCACAGTGTCCAGTGGGATCAGCACAGTCCACCAATTCTGAGTCCAGCACACCTCC
 AGTGGGGCCAACACAGCCACCAACTCTGAGTCCAGTACGACCTCCAGTGGGGCCAACACAGCCAC
 CAACTCTGAGTCCAGCACAGTGTCCAGTGGGGCCAGCACTGCCACCAACTCTGAGTCCAGCACAA
 CCTCCAGTGGGGTCAGCACAGCCACCAACTCTGAGTCCAGCACAACTCCAGTGGGGCTAGCACA
 GCCACCAACTCTGACTCCAGCACAACTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCTAG
 CACAGTGTCCAGTGGGATCAGCACAGTCCACCAATTCTGAGTCCAGCACAACTCCAGTGGGGCCA
 ACACAGCCACCAACTCTGGGTCCAGTGTGACCTCTGCAGGCTCTGGAACAGCAGCTCTGACTGGA
 ATGCACACAACTTCCCATAGTGCATCTACTGCAGTGAAGTGAAGCAAAGCCTGGTGGGTCCCTGGT
 GCCGTGGGAAATCTTCTCATCACCTGGTCTCGGTTGTGGCGGCCGTGGGGCTCTTTGCTGGGC
 TCTTCTTCTGTGTGAGAAACAGCCTGTCCCTGAGAAACACCTTTAACACAGCTGTCTACCACCT
 CATGGCCTCAACCATGGCCTTGGTCCAGGCCCTGGAGGGAATCATGGAGCCCCCACAGGCCAG
 GTGGAGTCTAAGTGGTCTGGAGGAGACAGTATCATCGATAGCCATGGAGATGAGCGGGAGGA
 ACAGCGGGCCCTGAGCAGCCCCGGAAGCAAGTGCCGATTCTTCAGGAAGGAAGAGACCTGGGCA
 CCCAAGACCTGGTTTCTTTTATTTCATCCCAGGAGACCCCTCCCAGCTTTGTTTGGATCCTGAA
 AATCTTGAAGAAGGTATTCCTCACCTTTCTTGCCTTTACCAGACACTGGAAAGAGAATACTATAT
 TGCTCATTTAGCTAAGAAATAAATACATCTCATCTAACACACAGCAAAGAGAAGCTGTGCTTG
 CCCCAGGGTGGGTATCTAGCTCTGAGATGAATCAGTTATAGGAGAAAACCTCCATGCTGGACTC
 CATCTGGCATTCAAAATCTCCACAGTAAATCCAAAGACCTCAAAAAAAAAAAAAAAAAAAAAA
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 100

MKMQKGNVLLMFGLLLHLEAATNSNETSTSANTGSSVISSGASTATNSGSSVTSSGVSTATISGS
SVTSNGVSIVTNSEFHTTSSGISTATNSEFSTASSGISIATNSESSTTSSGASTATNSESSTPSS
GASTVTNSGSSVTSSGASTATNSESSTVSSRASTATNSESSTLSSGASTATNSDSSTTSSGASTA
TNSESSTTSSGASTATNSESSTVSSRASTATNSESSTTSSGASTATNSESRTTSNGAGTATNSES
STTSSGASTATNSDSSTVSSGASTATNSESSTTSSGASTATNSESSTTSSGASTATNSDSSTTSS
GAGTATNSESSTVSSGISTVTNSESSTPSSGANTATNSESSTTSSGANTATNSESSTVSSGASTA
TNSESSTTSSGVSTATNSESSTTSSGASTATNSDSSTTSSEASTATNSESSTVSSGISTVTNSES
STTSSGANTATNSGSSVTSAGSGTAALTGMHTTSHSASTAVSEAKPGGSLVPWEIFLITLVSVVA
AVGLFAGLFFCVRNSLSLRNTFNTAVYHPHGLNHGLGPGPGGNHGAPHRPRWSPNWFWRREPSSI
AMEMSGRNSGP

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 510-532

FIGURE 101

GGCCGGACGCCTCCGCGTTACGGGATGAATTAACGGCGGGTTCCGCACGGAGGTTGTGACCCCTA
CGGAGCCCCAGCTTGCCACGCACCCCACTCGGCGTCGCGCGGCGTGCCCTGCTTGTACAGGTG
GGAGGCTGGAATATCAGGCTGAAAAACAGAGTGGGTACTCTCTTCTGGGAAGCTGGCAACAAAT
GGATGATGTGATATATGCATTCCAGGGGAAGGAAATGTGGTGCTTCTGAACCCATGGTCAATT
AACGAGGCAGTTTCTAGCTACTGCACGTACTTCATAAAGCAGGACTCTAAAAGCTTTGGAATCAT
GGTGTGATGGAAGGGATTACTTTTACTGACTCTGTTTTGGGGAAGCTTTTTTGAAGCATTT
TCATGCTGAGTCCCTTTTTACCTTTGATGTTTGTAACCCATCTTGGTATCGCTGGATCAACAAC
CGCCTTGTGGCAACATGGCTCACCTACCTGTGGCATTATTGGAGACCATGTTTGGTGTAAAAGT
GATTATAACTGGGGATGCATTTGTTCTGGAGAAAGAAGTGTCAATTATCATGAACCATCGGACAA
GAATGGACTGGATGTTCTGTGGAATTGCCTGATGCGATATAGCTACCTCAGATTGGAGAAAATT
TGCCTCAAAGCGAGTCTCAAAGGTGTTCTGGATTGTTGGTTGGGCCATGCAGGCTGCTGCCTATAT
CTTCATTATAGGAAATGGAAGGATGACAAGAGCCATTTTGAAGACATGATTGATTACTTTTGTG
ATATTCACGAACCACTTCAACTCCTCATATTCCAGAAGGGACTGATCTCACAGAAAACAGCAAG
TCTCGAAGTAATGCATTTGCTGAAAAAATGGACTTCAGAAATATGAATATGTTTTACATCCAAG
AACTACAGGCTTTACTTTTGTGGTAGACCGTCTAAGAGAAGGTAAGAACCCTTGATGCTGTCCATG
ATATCACTGTGGCGTATCCTCACAACATTCTCAATCAGAGAAGCACCTCCTCCAAGGAGACTTT
CCCAGGGAAATCCACTTTCACGTCCACCGGTATCCAATAGACACCTCCCCACATCCAAGGAGGA
CCTTCAACTCTGGTGCCACAAACGGTGGGAAGAGAAAGAAGAGAGGCTGCGTTCTTCTATCAAG
GGGAGAAGAATTTTTATTTTACCGGACAGAGTGTCAATCCACCTGCAAGTCTGAACTCAGGGTC
CTTGTGGTCAAATTGCTCTCTATACTGTATTGGACCCTGTTTCAGCCCTGCAATGTGCCTACTCAT
ATATTTGTACAGTCTTGTAAAGTGGTATTTTATAATCACCATTGTAATCTTTGTGCTGCAAGAGA
GAATATTTGGTGGACTGGAGATCATAGAACTTGCATGTTACCGACTTTTACACAAACAGCCACAT
TTAAATTCAAAGAAAAATGAGTAAGATTATAAGGTTTGCCATGTGAAAACCTAGAGCATATTTG
GAAATGTTCTAAACCTTTCTAAGCTCAGATGCATTTTTCATGACTATGTCGAATATTTCTTACT
GCCATCATTATTTGTTAAAGATATTTTGCACCTAATTTTGTGGGAAAAATATTGCTACAATTTTT
TTAATCTCTGAATGTAATTCGATACTGTGTACATAGCAGGGAGTGATCGGGGTGAAATAACTT
GGGCCAGAATATTATTAAACAATCATCAGGCTTTTAAA

FIGURE 102

MHSRGREIVVLLNPWSINEAVSSYCTYFIKQDSKSGIMVSWKGIYFILTLFWGSFFGSIFMLSP
FLPLMFVNPSWYRWINNRLVATWLTLPVALLETMFGVKVIITGDAFVPGERSVIIMNHRTRMDWM
FLWNCLMRYSYLRLEKICLKASLKGVPFGFWAMQAAAYIFIHRKWKDDKSHFEDMIDYFCDIHEP
LQLLIFPEGTDLTENSKSRSNFAEKNGLQKYEYVLHPRTTGFTFVVDRLREGKNLDAVHDITVA
YPHNIPQSEKHLQGDFFPREIHFHVHRYPIDTLPTSKEDLQLWCHKRWEEKEERLRSFYQGEKNF
YFTGQSVIPPCKSELRLVLVVKLLSILYWTLFSPAMCLLIYLYSLVKWYFIITIVIFVLQERIFGG
LEIIELACYRLLHKQPHLNSKKNE

Important features of the protein:

Signal peptide:

amino acids 1-22

Transmembrane domains:

amino acids 44-63, 90-108, 354-377

FIGURE 103

CGGCTCGAGCGGCTCGAGTGAAGAGCCTCTCCACGGCTCCTGCGCCTGAGACAGCTGGCCTGACC
 TCCAAATCATCCATCCACCCCTGCTGTCATCTGTTTTTCATAGTGTGAGATCAACCCACAGGAATA
 TCCATGGCTTTTGTGCTCATTTTGGTTCTCAGTTTCTACGAGCTGGTGTGAGGACAGTGGCAAGT
 CACTGGACCGGGCAAGTTTGTCCAGGCTTGGTGGGGGAGGACGCCGTGTTCTCTGCTCCCTCT
 TTCCTGAGACCAAGTGCAGAGGCTATGGAAGTGCCTTCTTCAGGAATCAGTTCCATGCTGTGGTC
 CACCTCTACAGAGATGGGGAAGACTGGAATCTAAGCAGATGCCACAGTATCGAGGGAGAACTGA
 GTTTGTGAAGGACTCCATTGCAGGGGGCGTGTCTCTAAGGCTAAAAACATCACTCCCTCGG
 ACATCGGCCTGTATGGGTGCTGGTTCAGTTCCTCAGATTTACGATGAGGAGGCCACCTGGGAGCTG
 CGGGTGGCAGCACTGGGCTCACTTCTCTCATTTCATCGTGGGATATGTTGACGGAGGTATCCA
 GTTACTCTGCCTGTCTCAGGCTGGTTCCCCCAGCCACAGCCAAGTGGAAAGGTCCACAAGGAC
 AGGATTTGTCTTCAGACTCCAGAGCAATGCAGATGGGTACAGCCTGTATGATGTGGAGATCTCC
 ATTATGTCAGGAAAAATGCTGGGAGCATATTGTGTTCCATCCACCTTGCTGAGCAGAGTCATGA
 GGTGGAATCCAAGGTATTTGATAGGAGAGACGTTTTTCCAGCCCTCACCTTGGCGCTGGCTCTTA
 TTTTACTCGGGTTACTCTGTGGTGCCTGTGTGGTGTGTCTATGGGGATGATAATTGTTTTCTTC
 AAATCCAAAGGGAAAAATCCAGGCGGAAGTGGACTGGAGAAGAAAGCACGGACAGGCAGAATTGAG
 AGACGCCCGGAAACACGCAGTGGAGGTGACTCTGGATCCAGAGACGGCTCACCCGAAGCTGTGG
 TTTCTGATCTGAAAAGTGAACCCATAGAAAAGCTCCCCAGGAGGTGCCTCACTCTGAGAAGAGA
 TTTACAAGGAAGAGTGTGGTGGCTTCTCAGGGTTTCCAAGCAGGGAGACATTACTGGGAGGTGGA
 CGTGGGACAAAAATGTAGGGTGGTATGTGGGAGTGTGTGGGATGACGTAGACAGGGGGAAGAACA
 ATGTGACTTTGTCTCCCAACAATGGGTATTGGGTCTCAGACTGACAACAGAACATTTGTATTTT
 ACATTCAATCCCCATTTTATCAGCCTCCCCCAGCACCCCTCCTACACGAGTAGGGGTCTTCCT
 GGACTATGAGGGTGGGACCATCTCTTCTTCAATACAAATGACCAGTCCCTTATTTATACCTGC
 TGACATGTCAGTTTGAAGGCTTGTGAGACCTATATCCAGCATGCGATGTATGACGAGGAAAAG
 GGGACTCCCATATTCATATGTCCAGTGTCTGGGGATGAGACAGAGAAGACCTGCTTAAAGGGC
 CCCACACCACAGACCCAGACACAGCCAAGGGAGAGTGTCTCCGACAGGTGGCCCCAGCTTCTCT
 CCGGAGCCTGCGCACAGAGAGTCACGCCCCCACTCTCTTTAGGGAGCTGAGGTTCTTCTGCCC
 TGAGCCCTGCAGCAGCGGCAGTCACAGCTTCCAGATGAGGGGGGATTGGCCTGACCCTGTGGGAG
 TCAGAAGCCATGGCTGCCCTGAAGTGGGACGGAATAGACTCACATTAGGTTTAGTTTGTGAAA
 CTCCATCCAGCTAAGCGATCTTGAACAAGTCACAACTCCCAGGCTCCTCATTGCTAGTCACGG
 ACAGTGATTCTGCCTCACAGGTGAAGATTAAAGAGACAACGAATGTGAATCATGCTTGCAGGTT
 TGAGGGCACAGTGTGCTAATGATGTGTTTTTATATTATACATTTCCACCATAAACTCTGTT
 TGCTTATTCACATTAATTTACTTTTCTCTATACCAAATCACCATGGAATAGTTATTGAACACC
 TGCTTTGTGAGGCTCAAAGAATAAAGAGGAGGTAGGATTTTTCACTGATTCTATAAGCCCAGCAT
 TACCTGATACCAAACAGGCAAAGAAAACAGAAGAAGAGGAAGGAAAACACAGGTCCATATCC
 CTCATTAACACAGACACAAAATTTAAATAAAATTTTAACAAATTAACTAAACAATATATTTA
 AAGATGATATATAACTACTCAGTGTGGTTGTCCCAAAATGCAGAGTTGGTTTAATATTTAAAT
 ATCAACCAGTGAATTCAGCACATTAATAAAGTAAAAAGAAAACCATAAAAA

FIGURE 104

MAFVLILVLSFYELVSGQWQVTGPGKFVQALVGEDAVFSCSLFPETSAEAMEVRFFRNQFHAVVH
LYRDGEDWESKQMPQYRGRTEFVKDSIAGGRVSLRLKNITPSDIGLYGCWFSSQIYDEEATWELR
VAALGSLPLISIVGYVDGGIQLLCLSSGWFPQPTAKWKGPQGDLSSDSRANADGYSLYDVEISI
IVQENAGSILCSIHIAEQSHEVESKVLIGETFFQPSPWRLASILLGLLCGALCGVVMGMIIIVFFK
SKGKIQAELDWRKKGQAELRDARKHAVEVTLPETAHPKLCVSDLKTVTHRKAPQEVPHSEKRF
TRKSVVASQGFQAGRHYWEVDVGQNVGWYVGVCRDDVDRGKNNVTLSPNNGYWVLRLTTEHLYFT
FNPHFISLPPSTPPTRVGVFLDYEGGTISFFNTNDQSLIYTLTLCQFEGLLRPYIQHAMYDEEKG
TPIFICPVSWG

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 131-150, 235-259

FIGURE 105

CCTTCACAGGACTCTTCATTGCTGGTTGGCAATGATGTATCGGCCAGATGTGGTGAGGGCTAGGAAAAGAG
TTTGTGGGAACCCCTGGGTTATCGGCCCTCGTCATCTTCATATCCCTGATTGTCTGGCAGTGTGCATTGGA
CTCACTGTTTATTATGTGAGATATAATCAAAAGAAGACCTACAATTACTATAGCACATTGTCATTTACAAC
TGACAAACTATATGCTGAGTTTGGCAGAGAGGCTTCTAACAATTTTACAGAAATGAGCCAGAGACTTGAAT
CAATGGTGAAAAATGCATTTTATAAATCTCCATTAAGGGAAGAATTTGTCAAGTCTCAGGTTATCAAGTTC
AGTCAACAGAAGCATGGAGTGTGGCTCATATGCTGTTGATTGTAGATTTCACTCTACTGAGGATCCTGA
AACTGTAGATAAAATTGTTCAACTTGTTTTACATGAAAAGCTGCAAGATGCTGTAGGACCCCCCTAAAGTAG
ATCCTCACTCAGTTAAAAATTAATAAATCAACAAGACAGAAACAGACAGCTATCTAAACCATTGCTGCCGA
ACACGAAGAAGTAAACTCTAGGTGAGAGTCTCAGGATCGTTGGTGGGACAGAAGTAGAAGAGGGTGAATG
GCCCTGGCAGGCTAGCCTGCAGTGGGATGGGAGTCAATCGCTGTGGAGCAACCTTAATTAATGCCACATGGC
TTGTGAGTGCTGCTCACTGTTTACAACATATAAGAACCCTGCCAGATGGACTGCTTCCTTTGGAGTAACA
ATAAAACCTTCGAAAAATGAACGGGGTCTCCGGAGAATAATTGTCCATGAAAAATACAAACACCCATCACA
TGACTATGATATTTCTCTTGAGAGCTTTCTAGCCCTGTTCCCTACACAAATGCAGTACATAGAGTTTGTG
TCCCTGATGCATCCTATGAGTTTCAACCAGGTGATGTGATGTTTGTGACAGGATTTGGAGCACTGAAAAAT
GATGGTTACAGTCAAAATCATCTTCGACAAGCACAGGTGACTCTCATAGACGCTACAACCTTGCAATGAACC
TCAAGCTTACAATGAGCCATAACTCCTAGAATGTTATGTGCTGGCTCCTTAGAAGGAAAAACAGATGCAT
GCCAGGGTGACTCTGGAGGACCAGTGGTTAGTTCAGATGCTAGAGATATCTGGTACCTTGCTGGAATAGTG
AGCTGGGGAGATGAATGTGCGAAACCCAACAAGCCTGGTGTGTTATACTAGAGTTACGGCCTTGCGGGACTG
GATTACTTCAAAAACTGGTATCTTAAGAGACAAAAGCCTCATGGAACAGATAACATTTTTTTTTGTTTTTTG
GGTGTGGAGGCCATTTTATAGAGATACAGAATTGGAGAAGACTTGCAAAACAGCTAGATTTGACTGATCTCA
ATAAACTGTTTGCTTGATGCATGTATTTTCTTCCAGCTCTGTTCCGCACGTAAGCATCCTGCTTCTGCCA
GATCAACTCTGTCTATCTGTGAGCAATAGTTGAACTTTATGTACATAGAGAAATAGATAATACAATATTAC
ATTACAGCCTGTATTCATTTGTTCTCTAGAAGTTTGTGTCAGAAATTTGACTTGTGACATAAATTTGTAAT
GCATATATACAATTTGAAGCACTCCTTTCTTCAGTTCCTCAGCTCCTCTCATTTAGCAAATATCCATTT
TCAAGGTGCAGAACAGGAGTGAAAGAAAATATAAGAAGAAAAAATCCCCTACATTTTATTGGCACAGAA
AAGTATTAGGTGTTTTCTTAGTGGAATATTAGAAATGATCATATTCATTATGAAAGGTCAAGCAAAGACA
GCAGAATACCAATCACTTCATCATTTAGGAAGTATGGGAACTAAGTTAAGGAAGTCCAGAAAGAAGCCAAG
ATATATCCTTATTTTCAATTTCCAAACAACACTACTATGATAAATGTGAAGAAGATTCTGTTTTTTGTGACCT
ATAATAATTATACAACTTCATGCAATGTACTTGTCTAAGCAAATTAAGCAAATATTTATTTAACATTG
TTACTGAGGATGTCAACATATAACAATAAATATAAATCACCCA

FIGURE 106

MMYRPDVVRARKRVCWEPWVIGLVIFISLIVLAVCIGLTVHYVRYNQKKTNYNSTLSFTTDKLY
AEFGREASNNFTEMSQRLESMVKNAFYKSPLREEFVKSQVIKFSQQKHGVLAHMLLICRFHSTED
PETVDKIVQLVLHEKLQDAVGPPKVDPHSVKIKKINKTETDSYLNHCCGTRRSKTLGQSLRIVGG
TEVEEGEWPWQASLQWDGSHRCGATLINATWLVSAAHCFTTYKNPARWTASFGVTIKPSKMKRGL
RRIIVHEKYKHPSHDYDISLAELSSPVYTNVHRVCLPDASYEFQPGDVMFVTGFGALKNDGYS
QNHLRQAQVTLIDATTCNEPQAYNDAITPRMLCAGSLEGKTDACQGDGGPLVSSDARDIWYLAG
IVSWGDECAKPNKPGVYTRVTALRDWITSKTGI

Transmembrane domain:

amino acids 21-40 (type II)

FIGURE 107

AGAGAAAGAAGCGTCTCCAGCTGAAGCCAATGCAGCCCTCCGGCTCTCCGCGAAGAAGTTCCCTG
CCCCGATGAGCCCCCGCGTCCCGACTATCCCCAGGCGGGCGTGGGGCACCAGGGCCAGC
GCCGACGATCGCTGCCGTTTTGCCCTTGGGAGTAGGATGTGGTGAAAGGATGGGGCTTCTCCCTT
ACGGGGCTCACAATGGCCAGAGAAGATTCCGTGAAGTGTCTGCGCTGCCTGCTCTACGCCCTCAA
TCTGCTCTTTTGGTTAATGTCCATCAGTGTGTGGCAGTTTCTGCTTGGATGAGGGACTACCTAA
ATAATGTTCTCACTTTAACTGCAGAAACGAGGGTAGAGGAAGCAGTCATTTGACTTACTTTCCCT
GTGGTTCATCCGGTCATGATTGCTGTTTGTGCTTTTCTTATCATTGTGGGGATGTTAGGATATTG
TGGAACGGTGAAAAGAAATCTGTTGCTTCTTGCATGGTACTTTGGAAGTTTGTCTGCTATTTCT
GTGTAGAACTGGCTTGTGGCGTTTGGACATATGAACAGGAACCTTATGGTTCAGTACAATGGTCA
GATATGGTCACTTTGAAAGCCAGGATGACAAATATGGATTACCTAGATATCGGTGGCTTACTCA
TGCTTGAATTTTTTTCAGAGAGAGTTTAAAGTGTGTGGAGTAGTATATTTCACTGACTGGTTGG
AAATGACAGAGATGGACTGGCCCCAGATTCCCTGCTGTGTAGAGAATTTCCAGGATGTTCCAAA
CAGGCCACCAGGAAGATCTCAGTGACCTTTATCAAGAGGGTGTGGGAAGAAAATGTATTTCCTT
TTTGAGAGGAACCAACAACACTGCAGGTGCTGAGGTTTCTGGGAATCTCCATTGGGGTGACACAAA
TCCTGGCCATGATTCTCACCATTACTCTGCTCTGGGCTCTGTATTATGATAGAAGGGAGCCTGGG
ACAGACCAAATGATGTCCTTGAAGAATGACAACCTCAGCACCTGTCTGTCCTCAGTAGAACT
GTTGAAACCAAGCCTGTCAAGAATCTTTGAACACACATCCATGGCAAACAGCTTAAATACACACT
TTGAGATGGAGGAGTTATTAAGAAAGAAATGTACAGAAAGAAACCACAACTTGTTTTATTGGACT
TGTGAATTTTGGAGTACATATCTGTGTTTTCAGAAATATGTAGAAATAAAAATGTTGCCATAAAA
TAACACCTAAGCATATACTATTCTATGCTTTAAATGAGGATGGAAAAGTTTCATGTCATAAGTC
ACCACCTGGACAATAATTGATGCCCTTAAATGCTGAAGACAGATGTCATACCCACTGTGTAGCC
TGTGTATGACTTTTACTGAACACAGTTATGTTTTGAGGCAGCATGGTTTGATTAGCATTTCCGCA
TCCATGCAAACGAGTCACATATGGTGGGACTGGAGCCATAGTAAAGGTTGATTACTTCTACCAA
CTAGTATATAAAGTACTAATTAATGCTAACATAGGAAGTTAGAAAATACTAATAACTTTTATTA
CTCAGCGATCTATTCTTCTGATGCTAAATAAATTATATATCAGAAAACCTTCAATATTGGTGACT
ACCTAAATGTGATTTTTGCTGGTTACTAAATATTTCTTACCCTTAAAGAGCAAGCTAACACAT
TGTCTTAAGCTGATCAGGGATTTTTGTATATAAGTCTGTGTTAAATCTGTATAATTCAGTCGAT
TTCAGTTCTGATAATGTTAAGAATAACCATTATGAAAAGGAAAATTTGTCCTGTATAGCATCATT
ATTTTTAGCCTTTTCTGTTAATAAAGCTTTACTATTCTGTCTGGGCTTATATTACACATATAAC
TGTTATTTAAATACTTTAACCCTAATTTTGAAAATTACAGTGTGATACATAGGAATCATTATTC
AGAATGTAGTCTGGTCTTTAGGAAGTATTAATAAGAAAATTTGCACATAACTTAGTTGATTGAGA
AAGGACTTGATGCTGTTTTTCTCCAAATGAAGACTCTTTTGGACACTAAACACTTTTTAAAAA
GCTTATCTTTGCCTTCTCCAAACAAGAAGCAATAGTCTCCAAGTCAATATAAATCTACAGAAAA
TAGTGTCTTTTTCTCCAGAAAAATGCTTGTGAGAATCATTAAACATGTGACAATTTAGAGATT
CTTTGTTTTATTCTGATTAATATACTGTGGCAAATTACACAGATTATTAAATTTTTTTACAA
GAGTATAGTATATTTATTTGAAATGGGAAAAGTGCATTTTACTGTATTTTGTGATTTTGTGTTAT
TTCTCAGAATATGGAAGAAAAATTAATGTGTCAATAAATATTTCTAGAGAGTAA

FIGURE 108

MAREDSVKCLRCLLYALNLLFWLMSISVLAVSAWMRDYLNNVLTTLTAETRVEEAVILTYFPVVHP
VMIAVCCFLIIVGMLGYCGTVKRNLLLLAWYFGSLLVIFCVELACGVWTYEQELMVPVQWSDMT
LKARMTNYGLPRYRWLTHAWNFFQREFKCCGVVYFTDWLEMTMDWPPDSCCVREFPGCSKQAHQ
EDLSDLYQEGCGKKMYSFLRGTKQLQVLRFLGISIGVTQILAMILTITLLWALYYDRREPGTDQM
MSLKNDNSQHLSCPSVELLKPSLSRIFEHTSMANSFNTHFEMEEL

Signal peptide:

amino acids 1-33

Transmembrane domains:

amino acids 12-35, 57-86, 94-114, 226-248

FIGURE 109

CCAAGGCCAGAGCTGTGGACACCTTATCCCACCTCATCCTCATCCTCTTCTCTGATAAAGCCCCCTACCAGTGCT
GATAAAGTCTTTCTCGTGAGAGCCTAGAGGCCCTAAAAAAAAGTGCTTGAAAGAGAAGGGGACAAAGGAACA
CCAGTATTAAGAGGATTTTCCAGTGTTTCTGGCAGTTGGTCCAGAAGGATGCTCCATTCTGCTTCTCACCTG
CCTCTTCATCAGGCACCTCCGTGTACCCGTGGCCCTAGATCCTTGTCTGCTTACATCAGCCTGAATGAGC
CCTGGAGGAACACTGACCACCAGTTGGATGAGTCTCAAGGTCCTCCTCTATGTGACAACCATGTGAATGGGGAG
TGGTACCACTTCACGGGCATGGCGGGAGATGCCATGCCTACCTTCTGCATACCAGAAAACCACTGTGGAACCCA
CGCACCTGTCTGGCTCAATGGCAGCCACCCCTAGAAGGCGACGGCATTGTGCAACGCCAGGCTTGTGCCAGCT
TCAATGGGAACCTGCTGTCTCTGGAACACCACGCTGGAAGTCAAGGCTTGCCTTGGAGGCTACTATGTGTATCGT
CTGACCAAGCCCAGCGTCTGCTTCCACGTCTACTGTGGTCATTTTTATGACATCTGCGACGAGGACTGCCATGG
CAGCTGCTCAGATACCAGCGAGTGCACATGCGCTCCAGGAAGTGTGCTAGGCCCTGACAGGCAGACATGCTTTG
ATGAAAATGAATGTGAGCAAAACAACGGTGGCTGCAGTGAGATCTGTGTGAACCTCAAAAACCTCTACCCTGT
GAGTGTGGGGTGGCCGTGTGCTAAGAAGTGATGGCAAGACTTGTGAAGACGTTGAAGGATGCCACAATAACAA
TGCTGGCTGCAGCCACTCTTGCCTTGGATCTGAGAAAGGCTACCAGTGTGAATGTCCCCGGGGCCTGGTGTGT
CTGAGGATAACCACACTTGCCAAAGTCCCTGTGTGTGCAAATCAAATGCCATTGAAGTGAACATCCCCAGGGAG
CTGGTGGTGGCCTGGAGCTTCTCTGACCAACACCTCCTGCCGAGGAGTGTCCAACGGCACCCATGTCAACAT
CCTCTTCTCTCAAGACATGTGGTACAGTGGTCCATGTGGTGAATGACAAGATTGTGGCCAGCAACCTCGTGA
CAGGTCTACCCAAGCAGACCCCGGGGAGCAGCGGGGACTTCATCATCCGAACCAGCAAGCTGCTGATCCCGGTG
ACCTGCGAGTTTCCACGCCCTGTACACCATTCTGAAGGATACGTTCCCAACCTTCGAACTCCCCACTGGAAAT
CATGAGCCGAAATCATGGGATCTTCCCATTCACTCTGGAGATCTTCAAGGACAATGAGTTTGAAGAGCCTTACC
GGGAAGCTCTGCCACCCTCAAGCTTCGTGACTCCCTCTACTTTGGCATTGAGCCCGTGGTGCACGTGAGCGGC
TTGGAAAGCTTGGTGGAGAGCTGCTTTGCCACCCCACTCCAAGATCGACGAGGTCCTGAAATACTACCTCAT
CCGGGATGGCTGTGTTTCAGATGACTCGGTAAAGCAGTACACATCCCCGGGATCACCTAGCAAAGCACTTCCAGG
TCCCTGTCTTCAAGTTTGTGGGCAAAGACCACAAGGAAGTGTCTGCACTGCCGGGTCTTGTCTGTGGAGTG
TTGGACGAGCGTTCCCGCTGTGCCAGGGTTGCCACCGGCGAATCCGTCGTGGGGCAGGAGGAGAGGACTCAGC
CGGTCTACAGGGCCAGACGCTAACAGGCGGCCCGATCCGCATCGACTGGGAGGACTAGTTCGTAGCCATACCTC
GAGTCCCCTGCATTGGACGGCTCTGCTCTTTGGAGCTTCTCCCCCACCGCCCTCTAAGAACATCTGCCAACAGC
TGCGTTCAGACTTCACACTGTGAGTTCAGACTCCCAGCACCAACTCACTCTGATTCTGGTCCATTCACTGGGCA
CAGGTACAGCACTGTGAACAATGTGGCCTGGGTGGGTTTCATCTTTCTAGGGTTGAAAATAAACTGTCCA
CCCAGAAAGACACTCACCCCATTTCCCTCATTTCTTCTACACTTAAATACCTCGTGATGGTGAATCAGAC
CACAAAATCAGAAGCTGGGTATAATATTTCAAGTTACAAAACCTAGAAAAATTAAACAGTTACTGAAATTATGA
CTTAAATACCAATGACTCCTTAAATATGTAAATTATAGTTATACCTTGAAATTCAATTCAAATGCAGACTAA
TTATAGGGAATTTGGAAGTGTATCAATAAACAGTATATAATTTT

FIGURE 110

MPPFLLLTCLFITGTSVSPVALDPCSAYISLNEPWRNTDHLQDESQGPPLCDNHVNGEWYHFTGMAGDAMP
TFCIPENHCGTHAPVWLNGSHPLEGDGIVQRQACASENGNCCLWNTTVEVKACPGGYVYRLTKPSVCFHV
YCGHFYDIDCEDCHGSCSDTSECTCAPGTVLGPDRQTCFDENECEQNNGGCSEICVNLKNSYRCECGVGRV
LRSDGKTCEDVEGCHNNNGGCSHSLGSEKGYQCECPRLVLSNHTCQVPVLCKSNAIEVNIIPRELVGG
LELFLTNTSCRGVSNNGTHVNLFLSLKTCGTVDVVDKIVASNLVTGLPKQTPGSSGDFIIRTSKLLIPVT
CEFPRLYTISEGYVPNLRNSPLEIMSRNHGIFPFTLEIFKDNEFEPEPYREALPTLKLKRLDSLYFGIEPVVHV
SGLESLVESCFAPTSKIDEVLKYYLIRDGCVSDDSVKQYTSRDHLAKHFQVPVFKFVGKDHKEVFLHCRV
LVCGVLDERSRCAQGCHRRMRGAGGEDSAGLQGQTITGGPIRIDWED

Important features of the protein:

Signal peptide:

amino acids 1-16

N-glycosylation sites.

amino acids 89-93, 116-120, 259-263, 291-295, 299-303

Tyrosine kinase phosphorylation sites.

amino acids 411-418, 443-451

N-myristoylation sites.

amino acids 226-232, 233-239, 240-246, 252-258, 296-302, 300-306,
522-528, 531-537

Aspartic acid and asparagine hydroxylation site.

amino acids 197-209

2P domain proteins.

amino acids 431-457

Calcium-binding EGF-like proteins.

amino acids 191-212, 232-253

FIGURE 111

GAGAGAGGCAGCAGCTTGCTCAGCGGACAAGGATGCTGGGCGTGAGGGACCAAGGCCTGCCCTGCACTCGG
 GCCTCCTCCAGCCAGTGCTGACCAGGGACTTCTGACCTGCTGGCCAGCCAGGACCTGTGTGGGGAGGCCCT
 CCTGCTGCCCTTGGGGTGACAATCTCAGCTCCAGGCTACAGGGAGACCGGGAGGATCACAGAGCCAGCATGT
 TACAGGATCCTGACAGTGATCAACCTCTGAACAGCCTCGATGTCAAACCCCTGCGCAAACCCCGTATCCCC
 ATGGAGACCTTCAGAAAGGTGGGGATCCCCATCATCATAGCACTACTGAGCCTGGCGAGTATCATCATTGT
 GGTGTCTCATCAAGGTGATTCTGGATAAATACTACTTCTCTGCGGGCAGCCTCTCCACTTCATCCCGA
 GGAAGCAGCTGTGTACGGAGAGCTGGACTGTCCCTTGGGGGAGGACGAGGAGCACTGTGTCAAGAGCTTC
 CCGGAAGGGCCTGCAGTGGCAGTCCGCCTCTCCAAGGACCGATCCACACTGCAGGTGCTGGACTCGGCCAC
 AGGGAAGTGGTTCTCTGCCTGTTTCGACAACCTTCACAGAAGCTCTCGCTGAGACAGCCTGTAGGCAGATGG
 GCTACAGCAGAGCTGTGGAGATTGGCCCAGACCAGGATCTGGATGTTGTTGAAATCACAGAAAACAGCCAG
 GAGCTTCGCATCGGGAACCAAGTGGGCCCTGTCTCTCAGGCTCCCTGGTCTCCCTGCACTGTCTTGCCCTG
 TGGGAAGAGCCTGAAGACCCCCGTGTGGTGGGTGGGGAGGAGCCCTCTGTGGATTCTTGGCCTTGGCAGG
 TCAGCATCCAGTACGACAAACAGCACGTCTGTGGAGGGAGCATCCTGGACCCCCACTGGGTCTCACGGCA
 GCCCCACTGCTTCAGGAAACATACCGATGTGTTCAACTGGAAGGTGCGGGCAGGCTCAGACAAACTGGGCAG
 CTTCCCATCCCTGGCTGTGGCCAAGATCATCATCATGAATTCAACCCCATGTACCCCAAAGACAATGACA
 TCGCCCTCATGAAGCTGCAGTTCCCACTCACTTTCTCAGGCACAGTCAGGCCCATCTGTCTGCCCTTCTTT
 GATGAGGAGCTCACTCCAGCCACCCCACTCTGGATCATTGGATGGGGCTTTACGAAGCAGAATGGAGGGAA
 GATGTCTGACATACTGCTGCAGGCGTCAGTCCAGGTCAATTGACAGCACACGGTGCAATGCAGACGATGCGT
 ACCAGGGGGGAAGTCACCGAGAAGATGATGTGTGCAGGCATCCCGGAAGGGGGTGTGGACACCTGCCAGGGT
 GACAGTGGTGGGCCCTGATGTACCAATCTGACCAGTGGCATGTGGTGGGCATCGTTAGCTGGGGCTATGG
 CTGCGGGGGCCCGAGCACCCAGGAGTATACACCAAGGTCTCAGCCTATCTCAACTGGATCTACAATGTCT
 GGAAGGCTGAGCTGTAATGCTGCTGCCCCTTTGAGTGTCTGGGAGCCGCTTCCTTCCTGCCCTGCCACCT
 GGGGATCCCCCAAAGTCAGACACAGAGCAAGAGTCCCTTGGGTACACCCCTCTGCCACAGCCTCAGCAT
 TTCTTGGAGCAGCAAAGGGCCTCAATTCTGTAAAGAGACCTCGCAGCCCAGAGGCGCCAGAGGAAGTCA
 GCAGCCCTAGCTCGGCCACACTTGGTGCTCCAGCATCCCAGGGAGAGACACAGCCCACTGAACAAGGTCT
 CAGGGGTATTGCTAAGCCAAGAAGGAACTTTCCACACTACTGAATGGAAGCAGGCTGTCTTGAAAAAGCC
 CAGATCACTGTGGGCTGGAGAGGAGAAGGAAAGGGTCTGCGCCAGCCCTGTCCGTCTTCACCCATCCCCAA
 GCCTACTAGAGCAAGAAACCAGTTGTAATATAAAATGCACTGCCCTACTGTTGGTATGACTACCGTTACCT
 ACTGTTGTCATTGTTATTACAGCTATGGCCACTATTATTAAAGAGCTGTGTAAACATCTCTGGCAAAAAAA
 AAAA

FIGURE 112

MLQDPDSDQPLNSLDVKPLRKPRIPMETFRKVGIPIIIALLSLASIIIVVLIKVILDKYYFLCG
QPLHFIPRKQLCDGELDCPLGEDEEHCVKSFPEGPAVAVRLSKDRSTLQVLDSATGNWFSACFDN
FTEALAEACRQMGYSRAVEIGPDQDLDVVEITENSQELMRNSSGPCLSGSLVSLHCLACGKSL
KTPRVVGGEASVDSWPWQVSIQYDKQHVC GGSILDPHWLTAHCFRKHTDVFNNKVRAGSDKL
GSFPSLAVAKIIIIIEFNPMYPKDNDIALMKLQFPLTFSGTVRPICLPFFDEELTPATPLWIIGWG
FTKQNGGKMSDILLQASVQVIDSTRCNADDAYQGEVTEKMMCAGIPEGGVDTCCQDSSGGPLMYQS
DQWHVVGIVSWGYGCGGPSTPGVYTKVSAYLNWIYNVWKAEL

Transmembrane domain:

amino acids 32-53 (typeII)

FIGURE 113

GGCTGGACTGGAACCTCCTGGTCCCAAGTGATCCACCCGCCTCAGCCTCCCAAGGTGCTGTGATTA
TAGGTGTAAGCCACCGTGTCTGGCCTCTGAACAACTTTTTTTCAGCAACTAAAAAGCCACAGGAGT
TGAAGTGTAGGATTCTGACTATGCTGTGGTGGCTAGTGCTCCTACTCCTACCTACATTAAATC
TGTTTTTTGTTCTCTTGTAACTAGCCTTTACCTTCCTAACACAGAGGATCTGTCACCTGTGGCTCT
GGCCCAAACCTGACCTTCACTCTGGAACGAGAACAGAGGTTTCTACCCACACCGTCCCCTCGAAG
CCGGGGACAGCCTCACCTTGCTGGCCTCTCGCTGGAGCAGTGCCTCACCAACTGTCTCACGTCT
GGAGGCACTGACTCGGGCAGTGCAGGTAGCTGAGCCTCTTGCTAGCTGCGGCTTTCAAGGTGGGC
CTTGCCCTGGCCGTAGAAGGGATTGCACAAGCCGAAGATTTATAGGCGATGGCTCCCACTGCCC
AGGCATCAGCCTTGCTGTAGTCAATCACTGCCCTGGGGCCAGGACGGGCGGTGGACACCTGTCTCA
GAAGCAGTGGGTGAGACATCACGCTGCCCCGCCATCTAACCTTTTCATGTCCTGCACATCACCTG
ATCCATGGGCTAATCTGAACTCTGTCCCAAGGAACCCAGAGCTTGAGTGAGCTGTGGCTCAGACC
CAGAAGGGGTCTGCTTAGACCACCTGGTTTTATGTGACAGGACTTGCATTCTCCTGGAACATGAGG
GAACGCCGGAGGAAAGCAAAGTGGCAGGGAAGGAAGTGTGCCAAATTATGGGTGAGAAAAGATG
GAGGTGTTGGGTTATCACAAAGGCATCGAGTCTCCTGCATTCACTGGACATGTGGGGGAAGGGCTG
CCGATGGCGCATGACACACTCGGGACTCACCTCTGGGGCCATCAGACAGCCGTTTCCGCCCGAT
CCACGTACCAGCTGCTGAAGGGCAACTGCAGGCCGATGCTCTCATCAGCCAGGCAGCAGCCAAAA
TCTGCGATCACAGCCAGGGGCAGCCGTCTGGGAAGGAGCAAGCAAAGTGACCATTTCTCCTCCC
CTCCTTCCCTCTGAGAGGCCCTCCTATGTCCCTACTAAAGCCACCAGCAAGACATAGCTGACAGG
GGCTAATGGCTCAGTGTGGCCCAGGAGGTGAGCAAGGCCTGAGAGCTGATCAGAAGGGCCTGCT
GTGCGAACACGGAATGCCTCCAGTAAGCACAGGCTGCAAAATCCCAGGCAGGCAAGGACTGTGTGG
CTCAATTTAAATCATGTTCTAGTAATTGGAGCTGTCCCCAAGACCAAAGGAGCTAGAGCTTGGTT
CAAATGATCTCCAAGGGCCCTTATACCCAGGAGACTTTGATTTGAATTTGAAACCCCAAATCCA
AACCTAAGAACCAGGTGCATTAAGAATCAGTTATTGCCGGGTGTGGTGGCCTGTAATGCCAACAT
TTTGGGAGGCCGAGGCGGGTAGATCACCTGAGGTGAGGAGTTCAAGACCAGCCTGGCCAACATGG
TGAAACCCCTGTCTCTACTAAAAATACAAAAAACTAGCCAGGCATGGTGGTGTGTGCCTGTATC
CCAGCTACTCGGGAGGCTGAGACAGGAGAATTACTTGAACCTGGGAGGTGAAGGAGGCTGAGACA
GGAGAATCACTTCAGCCTGAGCAACACAGCGAGACTCTGTCTCAGAAAAATAAAAAAGAATTA
TGTTTATTTGTAA

FIGURE 114

MLWLVLLLLPTLKSVCFLVTSLYLPNTEDLSLWLWPKPDLHSGTRTEVSTHTVPSKPGTASPC
WPLAGAVPSPTVSRLEALTRAVQVAEPLGSCGFQGGPCPGRRRD

Signal peptide:
amino acids 1-15

FIGURE 115

CAGCAGTGGTCTCTCAGTCCCTCTCAAAGCAAGGAAAGAGTACTGTGTGCTGAGAGACCATGGCAA
AGAATCCTCCAGAGAATTGTGAAGACTGTCACATTCTAAATGCAGAAGCTTTTAAATCCAAGAAA
ATATGTAAATCACTTAAGATTTGTGGACTGGTGTGTTGGTATCCTGGCCCTAACTCTAATTGTCTT
GTTTTGGGGGAGCAAGCACTTCTGGCCGGAGGTACCCAAAAAGCCTATGACATGGAGCACACTT
TCTACAGCAATGGAGAGAAGAAGAAGATTTACATGGAATTGATCCTGTGACCAGAACTGAAATA
TTCAGAAGCGGAAATGGCACTGATGAAACATTGGAAGTGCACGACTTTAAAAACGGATACACTGG
CATCTACTTCGTGGGTCTTCAAAAATGTTTTATCAAACTCAGATTAAAGTGATTCTGAATTTT
CTGAACCAGAAGAGGAAATAGATGAGAATGAAGAAATTACCACAACCTTTCTTTGAACAGTCAGTG
ATTTGGGTCCCAGCAGAAAAGCCTATTGAAAACCGAGATTTTCTTAAAAATTCCAAAATTCTGGA
GATTTGTGATAACGTGACCATGTATTGGATCAATCCCACTCTAATATCAGTTTCTGAGTTACAAG
ACTTTGAGGAGGAGGAGAAGATCTTCACTTTCCTGCCAACGAAAAAAGGGATTGAACAAAAT
GAACAGTGGGTGGTCCCTCAAGTGAAAGTAGAGAAGACCCGTCACGCCAGACAAGCAAGTGAGGA
AGAACTTCCAATAAATGACTATACTGAAAATGGAATAGAATTGATCCCATGCTGGATGAGAGAG
GTTATTGTTGTATTTACTGCCGTCGAGGCAACCGCTATTGCCGCCGCGTCTGTGAACCTTTACTA
GGCTACTACCCATATCCATACTGCTACCAAGGAGGACGAGTCATCTGTCGTGTCATCATGCCTTG
TAAGTGGTGGGTGGCCCGCATGCTGGGGAGGGTCTTAATAGGAGGTTTGAGCTCAAATGCTTAAAC
TGCTGGCAACATATAATAAATGCATGCTATTCAATGAATTTCTGCCTATGAGGCATCTGGCCCT
GGTAGCCAGCTCTCCAGAATTACTTGTAGGTAATTCCTCTCTTCATGTTCTAATAAACTTCTACA
TTATCACCAAAAAAAAAAAAAAAAAA

FIGURE 116

MAKNPPENCEDCHILNAEAFKSKKICKSLKICGLVFGILALTILVLFWGSKHFWPEVPPKKAYDME
HTFYSNGEKKKIYMEIDPVTRTEIFRSGNGTDETLEVHDFKNGYTGIIYFVGLQKCFIKTQIKVIP
EFSEPEEEIDENEEITTTFFEQSVIWPVPAEKPIENRDFLKNKILEICDNVTMYWINPTLISVSE
LQDFEEEGEDLHFPANEKKGIEQNEQWVVPQVKVEKTRHARQASEEELPINDYTENGIEFDPMLD
ERGYCCIIYCRGNRYCRRVCEPLLGYYPYPYCYQGGRVICRVIMPCNWWWVARMLGRV

Important features of the protein:

Signal peptide:

amino acids 1-40

Transmembrane domain:

amino acids 25-47 (type II)

N-glycosylation sites.

amino acids 94-97, 180-183

Glycosaminoglycan attachment sites.

amino acids 92-95, 70-73, 85-88, 133-136, 148-151, 192-195, 239-
242

N-myristoylation sites.

amino acids 33-38, 95-100, 116-121, 215-220, 272-277

Microbodies C-terminal targeting signal.

amino acids 315-317

Cytochrome c family heme-binding site signature.

amino acids 9-14

FIGURE 117

GAGCTCCCCCTCAGGAGCGCGTTAGCTTCACACCTTCGGCAGCAGGAGGGCGGCAGCTTCTCGCAGGCGGCA
GGGCGGGCGGGCCAGGATCATGTCCACCACCACATGCCAAGTGGTGGCGTTCTCTGTCCATCTGCGGGCT
GGCCGGCTGCATCGCGGCCACCGGGATGGACATGTGGAGCACCAGGACCTGTACGACAACCCGTCACCT
CCGTGTTCCAGTACGAAGGGCTCTGGAGGAGCTGCGTGAGGCAGAGTTCAGGCTTCACCGAATGCAGGCCC
TATTTACCATCTCTGGGACTTCCAGCCATGCTGCAGGCAGTGCAGGCCCTGATGATCGTAGGCATCGTCTCT
GGGTGCCATTGGCCTCTGTGATCCATCTTTGCCCTGAAATGCATCCGCATTGGCAGCATGGAGGACTCTG
CCAAAGCCAAACATGACACTGACCTCCGGGATCATGTTTCATTGTCTCAGGTCTTTGTGCAATTGCTGGAGTG
TCTGTGTTTGCCAACATGCTGGTGACTAAGTCTGGATGTCCACAGCTAACATGTACACCGGCATGGGTGG
GATGGTGCAGACTGTTTCAGACCAGGTACACATTTGGTGCGGCTCTGTTCTGTTGGGCTGGGTCTGAGGCCC
TCACACTAATTGGGGGTGTGATGATGTGCATCGCCTGCCGGGGCTGGCACCAGAAGAAACCAACTACAAA
GCCGTTTCTTATCATGCCTCAGGCCACAGTGTTCCTTACAAGCCTGGAGGCTTCAAGGCCAGCACTGGGCTT
TGGGTCCAACACCAAAACAAGAAGATATACGATGGAGGTGCCCGCACAGAGGACGAGGTACAATCTTATC
CTTCCAAGCAGACTATGTGTAAATGCTCTAAGACCTCTCAGCACGGGCGGAAGAACTCCCGGAGAGCTCA
CCCAAAAACAAGGAGATCCCATCTAGATTTCTTCTTCTTGTACTCACAGCTGGAAGTTAGAAAAGCCT
CGATTTTCATCTTTGGAGAGGCCAAATGGTCTTAGCCTCAGTCTCTGTCTCTAAATATTCACCATAAAACA
GCTGAGTTATTTATGAATTAGAGGCTATAGCTCACATTTTCAATCCTCTATTTCTTTTTTAAATATAACT
TTCTACTCTGATGAGAGAATGTGGTTTAAATCTCTCTCTCACATTTTGATGATTTAGACAGACTCCCCCTC
TTCCCTCTAGTCAATAAACCCATTGATGATCTATTTCCAGCTTATCCCCAAGAAACTTTTGAAAGGAAA
GAGTAGACCCAAAGATGTTATTTTCTGCTGTTTGAATTTTGTCTCCCCACCCCAACTTGGCTAGTAATAA
ACACTTACTGAAGAAGAAGCAATAAGAGAAAGATATTGTAATCTCTCCAGCCATGATCTCGGTTTTCTT
ACACTGTGATCTTAAAGTTACCAAACCAAGTCATTTTTCAGTTTGAGGCAACCAAACCTTTCTACTGCTG
TTGACATCTTCTTATTACAGCAACACCATTCTAGGAGTTTCCCTGAGCTCTCCACTGGAGTCTCTTTCTGT
CGCGGGTCAGAAATTGTCCCTAGATGAATGAGAAAATTATTTTTTTAATTTAAGTCCATAATATAGTTAA
AATAAATAATGTTTTAGTAAATGATACACTATCTCTGTGAAATAGCCTCACCCCTACATGTGGATAGAAG
GAAATGAAAAATAATTGCTTTGACATTGTCTATATGGTACTTTGTAAAGTCATGCTTAAGTACAAATTCC
ATGAAAAGCTCACACCTGTAATCCTAGCACCTTTGGGAGGCTGAGGAGGAAGGATCACTTGAGCCCAGAGT
TCGAGACTAGCCTGGGCAACATGGAGAAGCCCTGTCTCTACAAAATACAGAGAGAAAAATCAGCCAGTCA
TGGTGGCATAACCTGTAGTCCCAGCATTCGGGGAGGCTGAGGTGGGAGGATCACTTGAGCCCAGGGAGGT
TGGGGCTGCAGTGAGCCATGATCACACCACTGCACCTCCAGCCAGGTGACATAGCGAGATCCTGTCTAAAAA
AATAAAAAATAAATAATGGAACACAGCAAGTCTTAGGAAGTAGGTTAAACTAATTTCTTAA

FIGURE 118

MSTTTCQVVAFLLSILGLAGCIAATGMDMWSTQDLYDNPVTSVFQYEGLWRSCVRQSSGFTECRP
YFTILGLPAMLQAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLTSGIMFIVSGL
CAIAGVSVFANMLVTNFWMSTANMYTGMGGMVQTVQTRYTFGAALFVGWVAGGLTLIGGVMMCIA
CRGLAPEETNYKAVSYHASGHSVAYKPGGFKASTGFGSNTKNKKIYDGGARTEDEVQSYPSKHDY
V

Signal peptide:

amino acids 1-23

Transmembrane domains:

amino acids 81-100, 121-141, 173-194

FIGURE 119

GGAAAACTGTTCTTCTGTGGCACAGAGAACCCTGCTTCAAAGCAGAAGTAGCAGTTCGGAGTCC
AGCTGGCTAAACTCATCCAGAGGATAATGGCAACCCATGCCTTAGAAATCGCTGGGCTGTTTCTTG
GTGGTGTGGAATGGTGGGCACAGTGGCTGTCACTGTGATGCCTCAGTGGAGAGTGTCCGCCTTCATT
GAAACAACATCGTGGTTTTTGAAACTTCTGGGAAGGACTGTGGATGAATTGCGTGAGGCAGGCTAA
CATCAGGATGCAGTGCAAAATCTATGATTCCCTGCTGGCTCTTCTCCGGACCTACAGGCAGCCAGAG
GACTGATGTGTGCTGCTCCGTGATGTCCTTCTTGGCTTTCATGATGGCCATCCTTGGCATGAAATGC
ACCAGGTGCACGGGGGACAATGAGAAGGTGAAGGCTCACATTCTGCTGACGGCTGGAATCATCTTCAT
CATCACGGGCATGGTGGTGTCTATCCCTGTGAGCTGGGTTGCCAATGCCATCATCAGAGATTTCTATA
ACTCAATAGTGAATGTTGCCAAAAACGTGAGCTTGAGAAGCTCTCTACTTAGGATGGACCACGGCA
CTGGTGTCTGATTGTTGGAGGAGCTCTGTTCTGCTGCGTTTTTTGTTGCAACGAAAAGAGCAGTAGCTA
CAGATACTCGATACTTCCCATCGCACAAACCAAAAAAGTTATCACACCGGAAAGAAGTCACCGAGCG
TCTACTCCAGAAGTCAGTATGTGTAGTGTGTATGTTTTTTAACTTTACTATAAAGCCATGCAAATG
ACAAAAATCTATATTACTTTCTCAAAATGGACCCCAAGAACTTTGATTTACTGTTCTTAACTGCCT
AATCTTAATTACAGGAAGTGCATCAGCTATTTATGATTCTATAAGCTATTTACGAGAATGAGATA
TTAAACCAATGCTTTGATTGTTCTAGAAAGTATAGTAATTTGTTTTCTAAGGTGGTTCAAGCATCTA
CTCTTTTATCATTTACTTCAAAATGACATTGCTAAAGACTGCATTATTTTACTACTGTAATTTCTCC
ACGACATAGCATTATGTACATAGATGAGTGTAACATTTATATCTCACATAGAGACATGCTTATATGGT
TTTATTTAAATGAAATGCCAGTCCATTACACTGAATAAATAGAACTCAACTATTGCTTTTCAGGGAA
ATCATGGATAGGGTTGAAGAAGGTACTATTAATTGTTTAAAAACAGCTTAGGGATTAATGTCTCTCA
TTTATAATGAAGATTAATAATGAAGGCTTTAATCAGCATTGTAAAGGAAATTGAATGGCTTTCTGATAT
GCTGTTTTTTAGCCTAGGAGTTAGAAATCCTAACTTCTTTATCCTCTCTCCCAGAGGCTTTTTTTTT
CTTGTGTATTAAATTAACATTTTTTAAACGCAGATATTTTGTCAAGGGGCTTTCATTCAAACGCTT
TTCCAGGGCTATACTCAGAAGAAAGATAAAAGTGTGATCTAAGAAAAAGTGATGGTTTTAGGAAAGTG
AAAATATTTTTGTTTTGTATTTGAAGAAGAATGATGCATTTTGACAAGAAATCATATATGTATGGAT
ATATTTTAATAAGTATTTGAGTACAGACTTTGAGGTTTCATCAATATAAATAAAGAGCAGAAAAATA
TGTCTTGGTTTTTCATTTGCTTACCAAAAAACAACAACAAAAAAGTTGTCCTTTGAGAACTTCACCT
GCTCCTATGTGGGTACCTGAGTCAAAATGTCAATTTTTGTTCTGTGAAAAATAAATTTCTTCTGTGA
CCATTTCTGTTTAGTTTTTACTAAAATCTGTAAATACTGTATTTTTCTGTTTATTCCAAATTTGATGAA
ACTGACAATCCAATTTGAAAGTTTGTGTGACGCTGTGTCTAGCTTAAATGAATGTGTTCTATTGCTT
TATACATTTATATTAATAAATTTGTACATTTTTCTAATT

FIGURE 120

MATHALEIAGLFLGGVGMVGTVAVTVMPOWRVSAFIENNIVVFENFW EGLWMNCVRQANIRMQCK
IYDSL LALSPDLQAARGLMCAASVMSFLAFMAILGMKCTRCTGDNEKVKAHILLTAGIIFIITG
MVVLIPVSWVANAIIRDFYNSIVNVAQKRELGEALYLGWTTALVLIVGGALFCCVFCCNEKSSSY
RYSIPSHRTTQKSYHTGKKSPSVYSRSQYV

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 82-101, 118-145, 164-188

FIGURE 121

GGAGAGAGGCGCGGGGTGAAAGGCGCATTGATGCAGCCTGCGGCGGCCTCGGAGCGCGGCGGAG
CCAGACGCTGACCACGTTCTCTCTCGGTCTCCTCCGCCTCCAGCTCCGCGCTGCCCGGCAGCC
GGGAGCCATGCGACCCCAGGGCCCCGCGCCTCCCCGCAGCGGCTCCGCGGCCTCCTGCTGCTCC
TGCTGCTGCAGTGCCCCGCGCCGTCGAGCGCCTCTGAGATCCCCAAGGGGAAGCAAAGGCGCAG
CTCCGGCAGAGGGAGGTGGTGGACCTGTATAATGGAATGTGCTTACAAGGGCCAGCAGGAGTGCC
TGGTCGAGACGGGAGCCCTGGGGCCAATGTTATTCCGGGTACACCTGGGATCCCAGGTCGGGATG
GATTCAAAGGAGAAAAGGGGGAATGTCTGAGGGAAAGCTTGAGGAGTCCTGGACACCCAACCTAC
AAGCAGTGTTTCATGGAGTTCATTGAATTATGGCATAGATCTTGGGAAAATTGCGGAGTGATACATT
TACAAAGATGCGTTCAAATAGTGTCTAAGAGTTTTGTTTCAGTGGCTCACTTCGGCTAAAATGCA
GAAATGCATGCTGTCAGCGTTGGTATTTACATTCAATGGAGCTGAATGTTTCAGGACCTCTTCCC
ATTGAAGCTATAATTTATTTGGACCAAGGAAGCCCTGAAATGAATTCAACAATTAATATTCATCG
CACTTCTTCTGTGGAAGGACTTTGTGAAGGAATTGGTGTCTGGATTAGTGGATGTTGCTATCTGGG
TTGGCACTTGTTTCAGATTACCCAAAAGGAGATGCTTCTACTGGATGGAATTCAGTTTCTCGCATC
ATTATTGAAGAACTACCAAAATAAATGCTTTAATTTTCATTGCTACCTCTTTTTTTATTATGCC
TTGGAATGGTTCACTTAAATGACATTTTAAATAAGTTTATGTATACATCTGAATGAAAAGCAAAG
CTAAATATGTTTACAGACCAAAGTGTGATTTACACTGTTTTTAAATCTAGCATTATTCATTTTG
CTTCAATCAAAGTGGTTTCAATATTTTTTTTAGTTGGTTAGAATACTTCTTCATAGTCACATT
CTCTCAACCTATAATTTGGAATATTGTTGTGGTCTTTTGTCTTTTCTCTTAGTATAGCATTTTTTA
AAAAAATATAAAGCTACCAATCTTGTACAATTGTAAATGTTAAGAATTTTTTTTATATCTGT
TAAATAAAAATTATTTCCAACA

FIGURE 122

MRPQGPAASPQRLRGLLLLLLLQLPAPSSASEIPKGKQKAQLRQREVVDLYNGMCLQGPAQVPGR
DGSPGANVIPGTPGIPGRDGFKEKGECLRESFEESWTPNYKQCSWSSLNYGIDLGKIAECTFTK
MRSNSALRVLFSGSLRLKCRNACCQRWYFTFNGAECGGLPIEAIYLDQGSPEMNSTINIHRTS
SVEGLCEGIGAGLVDVAIWVGTCSDYPKGDASTGWNSVSRIIIIEELPK

Signal peptide:

amino acids 1-30

Transmembrane domain:

amino acids 195-217

FIGURE 123

GCTGAGCGTGTGCGCGGTACGGGGCTCTCCTGCCTTCTGGGCTCCAACGCAGCTCTGTGGCTGAA
CTGGGTGCTCATCACGGGAAGTCTGGGCTATGGAATACAGATGTGGCAGCTCAGGTAGCCCCAA
ATTGCCTGGAAGAATACATCATGTTTTTCGATAAGAAGAAATGTAGGATCCAGTTTTTTTTTTA
ACCGCCCCCTCCCCACCCCCAAAAAACTGTAAAGATGCAAAAACGTAATATCCATGAAGATCC
TATTACCTAGGAAGATTTTGATGTTTGTCTGCGAATGCGGTGTGGGATTTATTTGTTCTTGGAG
TGTTCTGCGTGGCTGGCAAAGAATAATGTTCCAAAATCGGTCCATCTCCCAAGGGGTCCAATTTT
TCTTCTGGGTGTGACGAGCCCTGACTCACTACAGTGCAGCTGACAGGGGCTGTATGCAACTG
GCCCCAAGCCAAAGCAAAGACCTAAGGACGACCTTTGAACAATACAAAGGATGGGTTTCAATG
TAATTAGGCTACTGAGCGGATCAGCTGTAGCACTGGTTATAGCCCCACTGTCTTACTGACAATG
CTTTCTTCTGCCGAACGAGGATGCCTAAGGGCTGTAGGTGTGAAGGCAAAATGGTATATGTGA
ATCTCAGAAATTACAGGAGATACCCCAAGTATATCTGCTGGTTGCTTAGGTTTGTCCCTTCGCT
ATAACAGCCTTCAAAAAGTAAAGTATAATCAATTTAAAGGGCTCAACCAGCTCACCTGGCTATAC
CTTGACCATTAACCATATCAGCAATATTGACGAAAATGCTTTTAAATGGAATACGCAGACTCAAAGA
GCTGATTCTTAGTTCCAATAGAATCTCCTATTTTCTTAACAATACCTTCAGACCTGTGACAAATT
TACGGAAGCTGGATCTGTCTTATAATCAGCTGCATCTCTGGGATCTGAACAGTTTCGGGGCTTG
CGGAAGCTGCTGAGTTTACATTTACGGTCTAATCCCTGAGAACCATCCCTGTGCGAATATTCCA
AGACTGCCGCAACCTGGAAGTTTTGGACCTGGGATATAACCGGATCCGAAGTTTAGCCAGGAATG
TCTTTGCTGGCATGATCAGACTCAAAGAAGTTCACCTGGAGCACAATCAATTTTCCAAGCTCAAC
CTGGCCCTTTTTCAGAGTTGGTCAGCCTTCAGAACCTTTACTTGCAGTGGGAATAAAATCAGTGT
CATAGGACAGACCATGTCTGGACCTGGAGCTCCTTACAAAGGCTTGATTTATCAGGCAATGAGA
TCGAAGCTTTCAGTGGAGCCAGTGTTCCTCAGTGTGTCCCGAATCTGCAGCGCCTCAACCTGGAT
TCCAACAAGCTCACATTTATTGGTCAAGAGATTTTGGATTCTTGATATCCCTCAATGACATCAG
TCTTGCTGGGAATATATGGGAATGCAGCAGAAATATTTGCTCCCTTGTAAGTGGCTGAAAAGTT
TTAAAGGCTCAAGGGAGAATACAATTATCTGTGCCAGTCCCAAAGAGCTGCAAGGAGTAAATGTG
ATCGATGCAGTGAAGAACTACAGCATCTGTGGCAAAGTACTACAGAGAGGTTTGATCTGGCCAG
GGCTCTCCCAAAGCCGACGTTTAAAGCCCAAGCTCCCCAGGCCGAAGCATGAGAGCAAACCCCTT
TGCCCCGACGGTGGGAGCCACAGAGCCCGGCCAGAGACCGATGCTGACGCCGAGCACATCTCT
TTCCATAAAATCATCGCGGGCAGCGTGGCGCTTTTCTGTCCGTGCTCGTCATCTCTGCTGGTTAT
CTACGTGTCATGGAAGCGGTACCTGCGAGCATGAAGCAGCTGCAGCAGCGCTCCCTCATGCGAA
GGCACAGGAAAAAGAAAAGACAGTCCCTAAAGCAAATGACTCCCAGCACCCAGGAATTTTATGTA
GATTATAAACCACCAACACGGAGACCGAGATGCTGCTGAATGGGACGGGACCTTGCACCTA
TAACAAATCGGGCTCCAGGGAGTGTGAGGTATCAACCATTGTGATAAAAAGAGCTCTTAAAGCT
GGGAAATAAGTGGTCTTTATTGAACTCTGGTGACTATCAAGGGAACGCGATGCCCCCCTCCCC
TTCCCTCTCCCTCTCACTTTGGTGGCAAGATCCTTCTTGTCCGTTTTAGTGCATTCAATAACT
GGTCATTTTCTCTCATACATAATCAACCCATTGAAATTTAAATACCACAATCAATGTGAAGCTT
GAACCTCCGGTTTAAATATAATACCTATTGTATAAGACCCTTTACTGATTCCATTAATGTCCCATTT
GTTTTAAGATAAACTTCTTTTCATAGGTAATAAAAAAAA

FIGURE 124

MGFNVIRLLSGSAVALVIAPT VLLTMLSSAERGCPKGCRCEGKMVYCESQKLQEIPSSISAGCLG
LSLRYNQLQKLKYNQFKGLNQLTWLYLDHNHISNIDENAFNGIRRLKELILSSNRISYFLNNTFR
PVTNLRNLDLSYNQLHSLGSEQFRGLRKLLSLHLRSNSLRTPVRI FQDCRNLELLDLGYNRIRS
LARNVFAGMIRLKEHLHLEHNQFSKLNALFPRLVSLQNLQWNKISVIGQTMSWTWSSLQRLDL
SGNEIEAFSGPSVFQCVPNLQRLNLDNKLTFIGQEILDSWISLNDISLAGNIWECSRNICSLVN
WLKSFKGLRENTIICASPKELGQVNVIDAVKNYSICGKSTTERFDLARALPKPTFKPKLPRPKHE
SKPPLPPTVGATEPGPETDADAEHISFHKIIAGSVALFLSVLVILLVIYVSWKRYPASMKQLQQR
SLMRRHRKKKRQSLKQMPSTQEFYVDYKPTNTETSEMLLNGTGPCITYNKSGSRECEV

Important features of the protein:

Signal peptide:

amino acids 1-33

Transmembrane domain:

amino acids 420-442

N-glycosylation sites.

amino acids 126-129, 357-360, 496-499, 504-507

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 465-468

Tyrosine kinase phosphorylation site.

amino acids 136-142

N-myristoylation sites.

amino acids 11-16, 33-38, 245-250, 332-337, 497-502, 507-512

FIGURE 125

CCGTTATCGTCTTGCGCTACTGCTGAATGTCCGTCCCGGAGGAGGAGGAGAGGCTTTTGCCGCTG
ACCCAGAGATGGCCCCGAGCGAGCAAATTCCTACTGTCCGGCTGCGCGGCTACCGTGGCCGAGCT
AGCAACCTTTCCCCTGGATCTCACAAAACTCGACTCCAAATGCAAGGAGAAGCAGCTCTTGCTC
GGTTGGGAGACGGTGCAAGAGAATCTGCCCCCTATAGGGGAATGGTGCGCACAGCCCTAGGGATC
ATTGAAGAGGAAGGCTTTCTAAAGCTTTGGCAAGGAGTGACACCCGCCATTACAGACACGTAGT
GTATTCGGAGGTGCAATGGTCACATATGAACATCTCCGAGAGGTTGTGTTTGGCAAAAGTGAAG
ATGAGCATTATCCCCCTTTGGAAATCAGTCATTGGAGGGATGATGGCTGGTGTATTGGCCAGTTT
TTAGCCAATCCAACCTGACCTAGTGAAGGTTGAGATGCAAATGGAAGGAAAAAGGAACTGGAAGG
AAAACCATTGCGATTTTCGTGGTGTACATCATGCATTTGCAAAAATCTTAGCTGAAGGAGGAATAC
GAGGGCTTTGGGCAGGCTGGGTACCCAATATACAAAGAGCAGCACTGGTGAATATGGGAGATTTA
ACCACCTATGATACAGTGAAACACTACTTGGTATTGAATACACCACTTGAGGACAATATCATGAC
TCACGGTTTATCAAGTTTATGTTCTGGACTGGTAGCTTCTATTCTGGGAACACCAGCCGATGTCA
TCAAAAGCAGAATAATGAATCAACCACGAGATAAACAAGGAAGGGGACTTTTGTATAAATCATCG
ACTGACTGCTTGATTGAGGCTGTTCAAGGTGAAGGATTCATGAGTCTATATAAAGGCTTTTACC
ATCTTGGCTGAGAATGACCCCTTGGTCAATGGTGTCTGGCTTACTTATGAAAAAATCAGAGAGA
TGAGTGGAGTCAGTCCATTTTAA

FIGURE 126

MSVPEEEERLLPLTQRWPRASKFLLSGCAATVAELATFPLDLTKTRLQMGEAALARLGDGARES
APYRGMVRTALGIIIEEGFLKLWQGVTPAIYRHVVYSGGRMVITYEHLREVVFVGKSEDEHYPLWKS
VIGGMMAGVIGQFLANPTDLVKVQMOMEGKRKLEGKPLRFRGVHHAFAKILAEGGIRGLWAGWVP
NIQRAALVNMCDLTITYDTVKHYLVLNTPLEDNIMTHGLSSLCSGLVASILGTPADVIKSRIMNQP
RDKQGRGLLYKSSTDCLIQAVQGEGFMSLYKGFLPSWLRMTPWSMVFWLTYEKIREMSGVSPF

Transmembrane domains:

amino acids 25-38, 130-147, 233-248

1000
900
800
700
600
500
400
300
200
100
0
100
200
300
400
500
600
700
800
900
1000

FIGURE 127

CGCGGATCGGACCCAAGCAGGTCGGCGGGCGGCGGAGAGCGGCCGGGCGTCAGCTCCTCGAC
CCCCGTGTCGGGCTAGTCCAGCGAGGCGGACGGGCGGCGTGGGCCCATGGCCAGGCCCGGCATGG
AGCGGTGGCGCGACCGGCTGGCGCTGGTGACGGGGGCTCGGGGGGCATCGGCGCGGCGGTGGCC
CGGGCCCTGGTCCAGCAGGGACTGAAGGTGGTGGGCTGCGCCCGCACTGTGGGCAACATCGAGGA
GCTGGCTGCTGAATGTAAGAGTGCAGGCTACCCCGGGACTTTGATCCCCTACAGATGTGACCTAT
CAAATGAAGAGGACATCCTCTCCATGTTCTCAGCTATCCGTTCTCAGCACAGCGGTGTAGACATC
TGCATCAACAATGCTGGCTTGGCCCGGCCTGACACCCTGCTCTCAGGCAGCACCAGTGGTTGGAA
GGACATGTTCAATGTGAACGTGCTGGCCCTCAGCATCTGCACACGGGAAGCCTACCAGTCCATGA
AGGAGCGGAATGTGGACGATGGGCACATCATTAACATCAATAGCATGTCTGGCCACCGAGTGTTA
CCCCTGTCTGTGACCACTTCTATAGTGCCACCAAGTATGCCGTCACCTGCGCTGACAGAGGGACT
GAGGCAAGAGCTTCGGGAGGCCAGACCCACATCCGAGCCACGTGCATCTCTCCAGGTGTGGTGG
AGACACAATTGCGCTTCAAACCTCCACGACAAGGACCCTGAGAAGGCAGCTGCCACCTATGAGCAA
ATGAAGTGTCTCAAACCCGAGGATGTGGCCGAGGCTGTTATCTACGTCCTCAGCACCCCGCACA
CATCCAGATTGGAGACATCCAGATGAGGCCACGGAGCAGGTGACCTAGTGACTGTGGGAGCTCC
TCCTTCCCTCCCCACCCTTCATGGCTTGCCCTGCTGCTGCTGATTTTAGGTGTTGATTTCTGGAT
CACGGGATACCACTTCCTGTCCACACCCCGACCAGGGGCTAGAAAATTTGTTTGAGATTTTATA
TCATCTTGTCAAATTGCTTCAGTTGTAAATGTGAAAAATGGGCTGGGGAAAGGAGGTGGTGTCCC
TAATTGTTTTACTTGTTAACTTGTTCTTGTTGCCCTGGGCACTTGGCCTTTGTCTGCTCTCAGTG
TCTTCCCTTTGACATGGGAAAGGAGTTGTGGCCAAAATCCCCATCTTCTTGCACCTCAACGTCTG
TGGCTCAGGGCTGGGGTGGCAGAGGGAGGCCTTCACCTTATATCTGTGTTGTTATCCAGGGCTCC
AGACTTCCTCCTCTGCCTGCCCCACTGCACCCTCTCCCCCTTATCTATCTCCTTCTCGGCTCCCC
AGCCCAGTCTTGGCTTCTTGTCCCTCCTGGGGTCATCCCTCCACTCTGACTCTGACTATGGCAG
CAGAACACCAGGGCCTGGCCAGTGGATTTTCATGGTGATCATTAAAAAAGAAAAATCGCAACCAA
AAAAAAAAA

FIGURE 128

MARPGMERWRDRLALVTGASGGIGAAVARALVQQGLKVVGCA RTVG NIEELAAECKSAGYPGTLI
PYRCDLSNEEDILSMFSAIRSQHSQVDICINNAGLARPD TLLSGSTSGWKDMFN VNV LALSICTR
EAYQSMKERNVDDGHIININSMGHRVLP LSVTHFY SATKYAVTALTEGLRQELREAQTHIRATC
ISPGV VETQFAFKLHDKDPEKAAATYEQMKCLKPEDVAEAVIYVLSTPAHIQIGDIQMRPTEQVT

Important features of the protein:

Signal peptide:

amino acids 1-17

N-myristoylation sites.

amino acids 18-24, 21-27, 22-28, 24-30, 40-46, 90-96, 109-115,
199-205

Short-chain alcohol dehydrogenase.

amino acids 30-42, 104-114

FIGURE 129

AAC'TTCTACATGGGGCTCCTGCTGCTGGTGCCTTTCCTCAGCCTCCTGCCGGTGGCCCTACACCAT
CATGTCCCTCCCACCCTCCTTTGACTGCGGGCCGTTTCAGGTGCAGAGTCTCAGTTGCCCGGGAGC
ACCTCCCCCTCCCGAGGCAGTCTGCTCAGAGGGCCTCGGCCCAGAATCCAGTTCTGGTTTCATGC
CAGCCTGTAAAAGGCCATGGAACCTTGGGTGAATCACCGATGCCATTTAAGAGGGTTTTCTGCCA
GGATGGAAATGTTAGGTCGTTCTGTGTCTGCGCTGTTTCATTTAGTAGCCACCAGCCACCTGTGG
CCGTTGAGTGCTTGAAATGAGGAACTGAGAAAATTAATTTCTCATGTATTTTCTCATTTATTTA
TTAATTTTTTAAGTATAGTTGTACATATTTGGGGGTACATGTGATATTTGGATACATGTATACAA
TATATAATGATCAAATCAGGGTAACTGGGATATCCATCACATCAAACATTTATTTTTTATCTTT
TTAGACAGAGTCTCACTCTGTCACCCAGGCTGGAGTGCAGTGGTGCCATCTCAGCTTACTGCAAC
CTCTGCCTGCCAGGTTCAAGCGATTCTCATGCCCTCCACCTCCCAAGTAGCTGGGACTACAGGCAT
GCACCACAATGCCCAACTAATTTTTGTATTTTGTAGTAGAGACGGGGTTTTGCCATGTTGCCCAGG
CTGGCCTTGAACCTCCTGGCCTCAAACAATCCACTTGCCTCGGCCTCCCAAAGTGTTATGATTACA
GGCGTGAGCCACCGTGCTGGCCTAAACATTTATCTTTTCTTTGTGTTGGGAACCTTGAAATTAT
ACAATGAATTATTGTTAACTGTCTCCTGCTGTGCTATGGAACACTGGGACTTCTTCCCTCT
ATCTAACTGTATATTTGTACCAGTTAACCAACCGTACTTCATCCCCACTCCTCTATCCTTCCC
AACCTCTGATCACCTCATTCTACTCTCTACCTCCATGAGATCCACTTTTTTAGCTCCCATGTG
AGTAAGAAAATGCAATATTTGTCTTTCTGTGCCTGGCTTATTTCACTTAACATAATGACTTCCTG
TTCCATCCATGTTGCTGCAAATGACAGGATTTCTGTTTAAATTTCAATTAAAAATAACCACACATG
GCAAAAA

FIGURE 130

MGLLLLVLFSLLPVAYTIMSLPPSFDCGPFRCRVSVAREHLPSRGSLLRGPRPRIPVLVSCQPV
KGGHTLGESPMFPKRVFCQDGNVRSFCVCAVHFSSHQPPVAVECLK

Important features of the protein:

Signal peptide:

amino acids 1-18

N-myristoylation site.

amino acids 86-92

Zinc carboxypeptidases, zinc-binding region 2 signature.

amino acids 68-79

FIGURE 131

TTCTGAAGTAACGGAAGCTACCTTGTATAAAGACCTCAACACTGCTGACCATGATCAGCGCAGCCTGGAGC
ATCTTCCTCATCGGGACTAAAATTGGGCTGTTCCCTTCAAGTAGCACCTCTATCAGTTATGGCTAAATCCTG
TCCATCTGTGTGTCGCTGCGATGCGGGTTTCATTTACTGTAATGATCGCTTCTGACATCCATTCCAACAG
GAATACCAGAGGATGCTACAACTCTCTACCTTCAGAACCAACAAATAAATAATGCTGGGATTCCCTTCAGAT
TTGAAAACTTGCTGAAAGTAGAAAGAATATACCTATACCACAACAGTTTAGATGAATTTCTACCAACCT
CCCAAAGTATGTAAAAGAGTTACATTTGCAAGAAAATAACATAAGGACTATCACTTATGATTCACTTTCAA
AAATTCCCTATCTGGAAGAATTACATTTAGATGACAACCTCTGTCTCTGCAGTTAGCATAGAAGAGGGAGCA
TTCCGAGACAGCAACTATCTCCGACTGCTTTTCCCTGTCCCATAATCACCTTAGCACAATTCCCTGGGGTTT
GCCCAGGACTATAGAAGAACTACGCTTGGATGATAATCGCATATCCACTATTTTCATCACCATCTCTTCAAG
GTCTCACTAGTCTAAAACGCTGGTTCTAGATGGAAACCTGTTGAACAATCATGGTTTAGGTGACAAAGTT
TTCTTCAACCTAGTTAATTTGACAGAGCTGTCCCTGGTGCGGAATTCCTGACTGCTGCACCAGTAAACCT
TCCAGGCACAAACCTGAGGAAGCTTTATCTTCAAGATAACCACATCAATCGGGTCCCCCAATGCTTTTT
CTTATCTAAGGCAGCTCTATCGACTGGATATGTCCAATAATAACCTAAGTAATTTACCTCAGGGTATCTTT
GATGATTTGGACAATATAACACAACCTGATTCTTCGCAACAATCCCTGGTATTGCGGGTGCAAGATGAAATG
GGTACGTGACTGGTTACAATCACTACCTGTGAAGGTCAACGTGCGTGGGCTCATGTGCCAAGCCCCAGAAA
AGGTTTCGTGGGATGGCTATTAAGGATCTCAATGCAGAACTGTTTGATTGTAAGGACAGTGGGATTGTAAGC
ACCATTTCAGATAACCACTGCAATACCCAACACAGTGTATCTGCCCAAGGACAGTGGCCAGCTCCAGTGAC
CAACAGCCAGATATTAAAGAACCCCAAGCTCACTAAGGATCAACAAACCACAGGGAGTCCCTCAAGAAAAA
CAATTACAATTACTGTGAAGTCTGTACCTCTGATACCATTCAATCTCTTGGAACTTGCTCTACCTATG
ACTGCTTTTGAGACTCAGCTGGCTTAAACTGGGCCATAGCCCGGCATTTGGATCTATAACAGAAAAAATTTGT
AACAGGGGAACGCAGTGAGTACTTGGTCACAGCCCTGGAGCCTGATTACCCCTATAAAGTATGCATGGTTC
CCATGGAAACCAGCAACCTCTACCTATTTGATGAAACTCCTGTTTGTATTGAGACTGAAACTGCACCCCTT
CGAATGTACAACCCCTACAACCACCCCTCAATCGAGAGCAAGAGAAAGAACCTTACAAAAACCCCAATTTACC
TTTGCTGCCATCATTGGTGGGGCTGTGGCCCTGGTTACCATTGCCCTTCTTGCTTTAGTGTGTTGGTATG
TTCATAGGAATGGATCGCTCTTCTCAAGGAACCTGTGCATATAGCAAAGGGAGGAGAAGAAAGGATGACTAT
GCAGAAGCTGGCATAAGAAGGACAACCTCTATCCTGGAAATCAGGGAACTTCTTTTCAGATGTTACCAAT
AAGCAATGAACCCATCTCGAAGGAGGAGTTTGTAAATACACACCATATTTCTCCTAATGGAATGAATCTGT
ACAAAAAATCAGTGAAAGCAGTAGTAACCGAAGCTACAGAGACAGTGGTATTCAGACTCAGATCAC
TCACACTCATGATGCTGAAGGACTCACAGCAGACTTGTGTTTTGGGTTTTTTAAACCTAAGGGAGGTGATG
GT

FIGURE 132

MISAAWSIFLIGTKIGLFLQVAPLSVMAKSCPSVCRCDAGFIYCNDRFLTSIPTGIPEDATTLYL
QNNQINNAGIPSDLKNNLLKVERIYLYHNSLDEFPTNLPKYVKELHLQENNIRTITYDSLSKIPYL
EELHLLDDNSVSAVSIEEGAFRDSNYLRLLFLSRNHLSTIPWGLPRTIEELRLDDNRISTISSPSL
QGLTSLKRLVLDGNLLNNHGLGDKVFFNLVNLTELSLVRNSLTAAPVNLPGTNLRKLYLQDNHIN
RVPPNAFSYLRQLYRLDMSNNNLSNLPQGIFDDLDNITQLILRNNPWYCGCKMKWVRDWLQSLPV
KVNVRGLMCQAPEKVRGMAIKDLNAELFDCKDSGIVSTIQITTAIPNTVYPAQGQWPAPVTKQPD
IKNPKLTKDQQTGSPSRKTITITIVKSVTSDTIHISWKLALPMTALRLSWLKLGHSPAFGSITET
IVTGERSEYLVTALEPDSYKVCMPMETSNLYLFEDETPVCIETETAPLRMYNPTTTLNREQEKE
PYKNPNLPLAAIIGGAVALVTIALLALVCWYVHRNGSLFSRNCAYSKGRRRKDDYAEAGTKKDNS
ILEIRETSFQMLPISNEFISKEEFVIHTIFPPNGMNLYKNNHSESSSNRSYRDSGIPDSHSHS

Important features of the protein:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 531-552

N-glycosylation sites.

amino acids 226-229, 282-285, 296-299, 555-558, 626-629, 633-636

Tyrosine kinase phosphorylation site.

amino acids 515-522

N-myristoylation sites.

amino acids 12-17, 172-177, 208-213, 359-364, 534-539, 556-561,
640-645

Amidation site.

amino acids 567-570

Leucine zipper pattern.

amino acids 159-180

Phospholipase A2 aspartic acid active site.

amino acids 34-44

FIGURE 133

CCGTCATCCCCCTGCAGCCACCCTTCCCAGAGTCCTTTGCCCAGGCCACCCCAGGCTTCTTGGCA
GCCCTGCCGGGCCACTTGTCTTCATGTCTGCCAGGGGGAGGTGGGAAGGAGGTGGGAGGAGGGCG
TGCAGAGGCAGTCTGGGCTTGGCCAGAGCTCAGGGTGCTGAGCGTGTGACCAGCAGTGAGCAGAG
GCCGGCCATGGCCAGCCTGGGGTGCTGCTCCTGCTCTTACTGACAGCACTGCCACCGCTGTGGT
CCTCCTCACTGCCCTGGGCTGGACACTGCTGAAAGTAAAGCCACCATTGCAGACCTGATCCTGTCT
GCGCTGGAGAGAGCCACCGTCTTCTAGAACAGAGGCTGCCTGAAATCAACCTGGATGGCATGGT
GGGGTCCGAGTGCTGGAAGAGCAGCTAAAAAGTGTCCGGGAGAAGTGGGCCACAGGAGCCCCCTGC
TGCAGCCGCTGAGCCTGCGCTGGGGATGCTGGGGGAGAAGCTGGAGGCTGCCATCCAGAGATCC
CTCCACTACCTCAAGCTGAGTGATCCCAAGTACCTAAGAGAGTTCCAGCTGACCCTCCAGCCCGG
GTTTTGGAAGCTCCACATGCCTGGATCCACACTGATGCCTCCTTGGTGTAACCCACGTTCCGGC
CCCAGGACTCATTCTCAGAGGAGAGAAGTGACGTGTGCCTGGTGCAGTGCTGGGAACCGGGACG
GACAGCAGCGAGCCCTGCGGCTCTCAGACCTCTGCAGGAGCCTCATGACCAAGCCCGGCTGCTC
AGGCTACTGCCTGTCCCACCAACTGCTCTTCTCTCTGCGCCAGAATGAGGGGATGCACACAGG
GACCACTCCAACAGAGCCAGGACTATATCAACCTCTTCTGCGCCAACATGATGGACTTGAACCGC
AGAGCTGAGGCCATCGGATACGCCTACCCTACCCGGGACATCTTCATGGAAAACATCATGTTCTG
TGGAATGGGCGGCTTCTCCGACTTCTACAAGCTCCGGTGGCTGGAGGCCATTCTCAGTGGCAGA
AACAGCAGGAAGGATGCTTCGGGAGCCTGATGCTGAAGATGAAGAATTATCTAAAGCTATTCAA
TATCAGCAGCATTTTTTCGAGGAGAGTGAAGAGGCGAGAAAAACAATTTCCAGATTCTCGCTCTGT
TGCTCAGGCTGGAGTACAGTGGCGCAATCTCGGCTCACTGCAACCTTTGCCTCCTGGGTTCAAGC
AATTCTCTTGCTCATCCTCCCGAGTAGCTGGGACTACAGGAGCGTGCCACCATACTGGCTAAT
TTTTATATTTTTTTAGTAGAGACAGGGTTTCATCATGTTGCTCATGCTGGTCTCGAACTCCTGAT
CTCAAGAGATCCGCCCACCTCAGGCTCCCAAAGTGTGGGATTATTAGGTGTGAGCCACCGTGTCTG
GCTGAAAAGCACTTTCAAAGAGACTGTGTTGAATAAAGGGCCAAGGTTCTTGCCACCCAGCACTC
ATGGGGGCTCTCTCCCTAGATGGCTGCTCCTCCACAACACAGCCACAGCAGTGGCAGCCCTGG
GTGGCTTCTATACATCCTGGCAGAATACCCCCAGCAAACAGAGAGCCACACCATCCACACCG
CCACCACCAAGCAGCCGCTGAGACGGACGGTTCCATGCCAGCTGCCTGGAGGAGGAACAGACCCC
TTTAGTCCTCATCCCTTAGATCCTGGAGGGCACGGATCACATCCTGGGAAGAAGGCATCTGGAGG
ATAAGCAAAGCCACCCCGACACCCAATCTTGAAGCCCTGAGTAGGCAGGGCCAGGGTAGGTGGG
GGCCGGGAGGGACCCAGGTGTGAACGGATGAATAAAGTTCAACTGCAACTGAAAAAAAAA

FIGURE 134

MSARGRWEGGRRACRGSIGLARAQGAERTVSSEQRPAMASLGLLLLLLLTALPPLWSSSLPGLD
TAESKATIADLILSALERATVFLEQRLPEINLDGMVGVRVLEEQLKSVREKWAQEPLLQPLSLRV
GMLGEKLEAAIQRSLSHYLKLSDPKYLREFQLTLPQGFWKLPHAWIHTDASLVYPTFGPQDSFSEE
RSDVCLVQLLGTGTDSEPCGLSDLCRSLMTKPGCSGYCLSHQLLFFLWARMRGCTQGPIQQSQD
YINLFCANMMDLNRRAEAIGYAYPTRDIFMENIMFCGMGGFSDFYKLRWLEAILSQKQEGCTG
EPDAEDEELSKAIQYQQHFSRRVKREKQFPDSRSVAQAGVQWRNLGSLQPLPPGFKQFSCILIP
SSWDYRSVPPYLANFYIFLVETGFHHVAHAGLELLISRDPPPTSGSQSVGL

Important features of the protein:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 39-56

Tyrosine kinase phosphorylation sites.

amino acids 149-156, 274-282

N-myristoylation sites.

amino acids 10-16, 20-26, 63-69, 208-214

Amidation site.

amino acids 10-14

Glycoprotein hormones beta chain signature 1.

amino acids 230-237

FIGURE 135

GGTCTGAGTGCAGAGCTGCTGTCTCATGGCGGCCGCTCTGTGGGGCTTCTTTCCCGTCCTGCTGCTG
CTGCTGCTATCGGGGGATGTCCAGAGCTCGGAGGTGCCCGGGGCTGCTGCTGAGGGATCGGGAGG
GAGTGGGGTCGGCATAGGAGATCGCTTCAAGATTGAGGGGCGTGCAAGTTGTTCCAGGGGTGAAGC
CTCAGGACTGGATCTCGGCGGCCCGAGTGTGGTAGACGGAGAAGAGCACGTCGGTTTCCTTAAG
ACAGATGGGAGTTTTGTGGTTCATGATATACCTTCTGGATCTTATGTAGTGAAGTTGTATCTCC
AGCTTACAGATTTGATCCCGTTTCGAGTGGATATCACTTCGAAAGGAAAAATGAGAGCAAGATATG
TGAATTACATCAAAACATCAGAGGTTGTTCAGACTGCCCTATCCTCTCCAAATGAAATCTTCAGGT
CCACCTTCTTACTTTTATTTAAAGGGGAATCGTGGGGCTGGACAGACTTTCTAATGAACCCAATGGT
TATGATGATGGTTCTTCCTTTATTGATATTTGTGCTTCTGCCCTAAAGTGGTCAACACAAGTGATC
CTGACATGAGACGGGAAATGGAGCAGTCAATGAATATGCTGAATTCCAACCATGAGTTGCCTGAT
GTTTCTGAGTTTCATGACAAGACTCTTCTCTTCAAATCATCTGGCAAATCTAGCAGCGGCAGCAG
TAAACAGGCAAAAGTGGGGCTGGCAAAGGAGGTAGTCAGGCCGTCCAGAGCTGGCATTTCAC
AAACACGGCAACACTGGGTGGCATCCAAGTCTTGGAACCGTGTGAAGCAACTACTATAAACTT
GAGTCATCCCGACGTTGATCTCTTACAACGTGTGTATGTT
AACTTTTTAGCACATGTTTTGTACTTGGTACACGAGAAAACCCAGCTTTCATCTTTTGTCTGTAT
GAGGTCAATATTGATGTCAGTGAATTAATTACAGTGTCTTATAGAAAATGCCATTAAATAATTAT
ATGAACTACTATACATTATGTATATTAATTAAACATCTTAATCCAGAAATCAAAAAAAAAAAAA
AAAAAAAAAAAAAA

FIGURE 136

MAAALWGFFPVLLLLLLSGDVQSSEVPGAAAEAGSGSGVGIGDRFKIEGRAVVPGVKPDWISAA
RVLVDGEEHVGLKTDGSFVVHDIPSGSYVVEVVS PAYRFD PVRVDITSKGKMRARYVNYIKTSE
VVRLPYPLQMKSSGPPSYFIKRESWGWTDFLMNPMVMMVLPLLI FVLLPKVVNTSDPDMRREME
QSMNMLNSNHELDPDVSEFMTRLFSSKSSGKSSSGSSKTGKSGAGKRR

Important features of the protein:

Signal sequence:

amino acids 1-23

Transmembrane domain:

amino acids 161-182

N-glycosylation site.

amino acids 184-187

Glycosaminoglycan attachment sites.

amino acids 37-40, 236-239

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 151-154

N-myristoylation sites.

amino acids 33-38, 36-41, 38-44, 229-234

Amidation site.

amino acids 238-241

ATP/GTP-binding site motif A (P-loop).

amino acids 229-236

FIGURE 137

GATGGCGCAGCCACAGCTTCTGTGAGATTTCGATTTCTCCCCAGTTCCCCTGTGGGTCTGAGGGGA
CCAGAAGGGTGAGCTACGTTGGCTTTCTGGAAGGGGAGGCTATATGCGTCAATTCCCCAAAACAA
GTTTTGACATTTCCCCTGAAATGTCATTCTCTATCTATTCACTGCAAGTGCCTGCTGTTCAGGC
CTTACCTGCTGGGCACTAACGGCGGAGCCAGGATGGGGACAGAATAAAGGAGCCACGACCTGTGC
CACCAACTCGCACTCAGACTCTGAACTCAGACCTGAAATCTTCTTTACGGGAGGCTTGGCAGT
TTTTCTTACTCCTGTGGTCTCCAGATTTCAAGCCTAAGATGAAAGCCTCTAGTCTTGCCTTCAGC
CTTCTCTCTGCTGCGTTTTATCTCCTATGGACTCCTTCCACTGGACTGAAGACACTCAATTTGGG
AAGCTGTGTGATCGCCACAAACCTTCAGGAAATACGAAATGGATTTTCTGAGATACGGGGCAGTG
TGCAAGCCAAAGATGGAACATTGACATCAGAATCTTAAGGAGGACTGAGTCTTTGCAAGACACA
AAGCCTGCGAATCGATGCTGCCTCCTGCGCCATTTGCTAAGACTCTATCTGGACAGGGTATTTAA
AACTACCAGACCCCTGACCATTATACTCTCCGGAAGATCAGCAGCCTCGCCAATTCCTTTCTTA
CCATCAAGAAGGACCTCCGGCTCTCTCATGCCACATGACATGCCATTGTGGGAGGAAGCAATG
AAGAAATACAGCCAGATTCTGAGTCACCTTGAAAAGCTGGAACCTCAGGCAGCAGTTGTGAAGGC
TTTGGGGGAACTAGACATTCTTCTGCAATGGATGGAGGAGACAGAATAGGAGGAAAGTGATGCTG
CTGCTAAGAATATTCGAGGTCAAGAGCTCCAGTCTTCAATACCTGCAGAGGAGGCATGACCCCAA
ACCACCATCTCTTTACTGTACTAGTCTTGTGCTGGTCACAGTGTATCTTATTTATGCATTACTTG
CTTCCTTGCAATGATTGTCTTTATGCATCCCCAATCTTAATTGAGACCATACTTGTATAAGATTTT
TGTAATATCTTTCTGCTATTGGATATATTTATTAGTTAATATATTTATTTATTTTGTCTATTTA
ATGTATTTATTTTTTACTTGGACATGAACTTTAAAAAATTACAGATTATATTTATAACCTG
ACTAGAGCAGGTGATGATTTTTTATACAGTAAAAAAAAAAACCTTGTAATTCTAGAAGAGTGG
CTAGGGGGGTATTCAATTTGTATTCAACTAAGGACATATTTACTCATGCTGATGCTCTGTGAGAT
ATTTGAAATTGAACCAATGACTACTTAGGATGGGTGTGGAATAAGTTTTGATGTGGAATTGCAC
ATCTACCTTACAATTACTGACCATCCCCAGTAGACTCCCCAGTCCCATAATTGTGTATCTTCCAG
CCAGGAATCCTACAGCCAGCATGTATTTCTACAAATAAAGTTTCTTTGCATACCAAAAAAAA
AAAAA

FIGURE 138

MRQFPKTSFDISPEMSFSIYSLQVPAVPGLTCWALTAEPGWGQNGATTTCATNSHSDSELRPEIF
SSREAWQFFLLWSPDFRPKMKASSLAFSLLSAAFYLLWTPSTGLKTLNLGSCVIATNLQEIRNG
FSEIRGSVQAKDGNIDIRILRRTESLQDTKPANRCCLLRHLLRLYLDRVFKNYQTPDHYTLRKIS
SLANSFLTICKDLRLSHAHMTCHCGEEAMKKYSQILSHFEKLEPQAAVVKALGELDILLQWMEET
E

Important features of the protein:

Signal peptide:

amino acids 1-42

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 192-195, 225-228

N-myristoylation sites.

amino acids 42-47, 46-51, 136-141

FIGURE 139

CCTGGAGCCGGAAGCGGGCTGCAGCAGGGCGAGGCTCCAGGTGGGGTCGGTTCCGCATCCAGCC
TAGCGTGTCCACGATGCGGGCTGGGCTCCGGGACTTTCGCTACCTGTTGCGTAGCGATCGAGGTGC
TAGGGATCGCGGTCTTCCTTCGGGGATTCTTCCCGGCTCCCGTTCGTTCTCTGCCAGAGCGGAA
CACGGAGCGGAGCCCCAGCGCCCGAACCCTCGGCTGGAGCCAGTTCTAACTGGACCACGCTGCC
ACCACCTCTCTTCAGTAAAGTTGTTATTGTTCTGATAGATGCCTTGAGAGATGATTTTGTGTTG
GGTCAAAGGGTGTGAAATTTATGCCCTACACAACCTTACCTTGTGGAAAAAGGAGCATCTCACAGT
TTTGTGGCTGAAGCAAAGCCACCTACAGTTACTATGCCTCGAATCAAGGCATTGATGACGGGGAG
CCTTCTCGGCTTTGTCGACGTCATCAGGAACCTCAATTCTCTGCACTGCTGGAAGACAGTGTGA
TAAGACAAGCAAAAGCAGCTGGAAAAAGAATAGTCTTTTATGGAGATGAAACCTGGGTAAATTA
TTCCCAAAGCATTGTGTGGAATATGATGGAACAACCTCATTTTCGTGTCAGATTACACAGAGGT
GGATAATAATGTCACGAGGCATTTGGATAAAGTATTAAAAAGAGGAGATTGGGACATATTAATCC
TCCACTACCTGGGGCTGGACCACATTGGCCACATTTAGGGCCCAACAGCCCCCTGATTGGGCAG
AAGCTGAGCGAGATGGACAGCGTGCTGATGAAGATCCACACCTCACTGCACTCGAAGGAGAGAGA
GACGCCCTTTACCCAATTTGCTGGTTCTTTGTGGTGACCATGGCATGTCTGAAACAGGAAGTCACG
GGGCTCTCCACCGAGGAGGTGAATACACCTCTGATTTTAAATCAGTTCTGCGTTTGAAGGAAA
CCCGGTGATATCCGACATCCAAAGCACGTCCAATAGACGGATGTGGCTGCGACACTGGCGATAGC
ACTTGGCTTACCGATTCCAAAAGACAGTGTAGGGAGCCTCCTATTCCCAGTTGTGGAAGGAAGAC
CAATGAGAGAGCAGTTGAGATTTTACATTTGAATACAGTGCAGCTTAGTAACTGTTGCAAGAG
AATGTGCCGTATATGAAAAAGATCCTGGGTTTGAGCAGTTTAAATGTGAGAAAGATTGCATGG
GAACTGGATCAGACTGTACTTGGAGGAAAAGCATTGAGAACTCCTATTCAACCTGGGCTCCAAGG
TTCTCAGGCAGTACCTGGATGCTCTGAAGACGCTGAGCTTGTCCCTGAGTGCACAAGTGGCCAG
TTCTCACCTCTGCTCTGCTCAGCGTCCCACAGGCACTGCACAGAAAGGCTGAGCTGGAAGTCCCA
CTGTCTCTCTCTGGGTTTTCTCTGCTCTTTTATTTGGTGATCCTGGTTCTTTCCGGCCGTTACAGT
CATTGTGTGCACCTCAGCTGAAAGTTCGTGCTACTTCTGTGGCCTCTCGTGGCTGGCGGCAGGCT
GCCTTTCGTTTACCAGACTCTGGTTGAACACCTGGTGTGTGCCAAGTGTGGCAGTGGCCTGGAC
AGGGGGCCTCAGGGAAGGACGTGGAGCAGCCTTATCCCAGGCTCTGGGTGTCCCGACACAGGTG
TTCACATCTGTGCTGTGAGGTGAGTGCCTCAGTTCTTGGAAAGCTAGGTTCCTGCGACTGTTAC
CAAGGTGATTGTAAAGAGCTGGCGGTACAGAGGAACAAGCCCCCAGCTGAGGGGGTGTGTGAA
TCGGACAGCCTCCAGCAGAGGTGTGGGAGCTGCAGCTGAGGGAAGAAGAGACAATCGGCCTGGA
CACTCAGGAGGGTCAAAAGGAGACTTGGTCGCCACCACTCATCCTGCCACCCCCAGAATGCATCCT
GCCTCATCAGGTCCAGATTTCTTTCCAAGGCGGACGTTTTCTGTTGGAATTCCTAGTCCTTGGCC
TCGGACACCTTCATTCTGTTAGCTGGGGAGTGGTGGTGAGGCAGTGAAGAAGAGGCGGATGGTCAC
ACTCAGATCCACAGAGCCCAGGATCAAGGGACCCACTGCAGTGGCAGCAGGACTGTTGGGCCCCC
ACCCCAACCCTGCACAGCCCTCATCCCCTCTTGGCTTGAAGCCGTGAGAGGCCCTGTGCTGAGTGT
CTGACCGAGACACTCACAGCTTTGTCATCAGGGCACAGGCTTCTCGGAGCCAGGATGATCTGTG
CCACGCTTGACCTCGGGCCCCATCTGGGCTCATGCTCTCTCTCTCTCTCTGCTATTGAATTAGTACCTAG
CTGCACACAGTATGTAGTTACCAAAAGAATAAACGGCAATAATTGAGAAAAAAA

FIGURE 140

MRLGSGTFATCCVAIEVLGIAVFLRGFFPAPVRSSARAETHGAEPPEPSAGASSNWTTLPPPLF
SKVVIVLIDALRDDDFVFGSKGVKEMPYTTYLVEKGASHSFVAEAKPPTVTMPRIKALMTGSLPGF
VDVIRNLNSPALLEDVIRQAKAAGKRIVFYGDETWVKLFPHFVEYDGTTSFFVSDYTEVDNNV
TRHLDKVLKRGDWDILILHYLGLDHHIGHISGPN SPLIGQKLSEMDSVLMKIHTSLQSKERETPLP
NLLVLCGDHGMSETGSHGASSTEEVNTPLILISSAFERKPGDIRHPKHVQ

Important features of the protein:

Signal peptide:

amino acids 1-34

Transmembrane domain:

amino acids 58-76

N-glycosylation sites.

amino acids 56-60, 194-198

N-myristoylation sites.

amino acids 6-12, 52-58, 100-106, 125-131, 233-239, 270-276,
275-281, 278-284

Amidation site.

amino acids 154-158

Cell attachment sequence.

amino acids 205-208

FIGURE 141

GGCACGAGGCAAGCCTTCCAGGTTATCGTGACGCACCTTGAAAGTCTGAGAGCTACTGCCCTACA
GAAAGTTACTAGTGCCCTAAAGCTGGCGCTGGCACTGATGTTACTGCTGCTGTTGGAGTACAAC
TCCCTATAGAAAACAACTGCCAGCACCTTAAGACCACTCACACCTTCAGAGTGAAGAACTTAAAC
CCGAAGAAATTCAGCATTCATGACCAGGATCACAAAGTACTGGTCCTGGACTCTGGGAATCTCAT
AGCAGTTCCAGATAAAAACTACATACGCCCAGAGATCTTCTTTGCATTAGCCTCATCCTTGAGCT
CAGCCTCTGCGGAGAAAGGAAGTCCGATTCTCCTGGGGGTCTCTAAAGGGAGTTTTGTCTCTAC
TGTGACAAGGATAAAGGACAAAGTCATCCATCCCTTCAGCTGAAGAAGGAGAACTGATGAAGCT
GGCTGCCCCAAAAGGAATCAGCACGCCGGCCCTTCATCTTTTATAGGGCTCAGGTGGGCTCCTGGA
ACATGCTGGAGTCGGCGGCTCACCCCGGATGGTTCATCTGCACCTCCTGCAATTGTAATGAGCCT
GTTGGGGTGACAGATAAATTTGAGAACAGGAAACACATTGAATTTTCATTTCAACCAGTTTGCAA
AGCTGAAATGAGCCCCAGTGAGGTCAGCGATTAGGAAACTGCCCCATTGAACGCCTTCTCGCTA
ATTTGAACTAATTGTATAAAAACACCAAACCTGCTCACT

FIGURE 142

MLLLLLLEYNFPIENNCQHLKTTHTFRVKNLNPKKFSIHDQDHKVLVLDSGNLIAVPDKNYIRPEI
FFALASSLSSASAEKGSPIILLGVSKGEFCLYCDKDKGQSHPSLQLKKEKLMKLAAQKESARRPFI
FYRAQVGSWNMLESAAHPGWFICTSCNCNEPVGVTDFENRKHIEFSFQPVCKAEMSPSEVSD

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 33-36

N-myristoylation site.

amino acids 50-55, 87-92

Interleukin-1

amino acids 37-182

FIGURE 143

CTAGAGAGTATAGGGCAGAAGGATGGCAGATGAGTGACTCCACATCCAGAGCTGCCTCCCTTTAA
TCCAGGATCCTGTCTTCTGTCTGTAGGAGTGCCTGTTGCCAGTGTGGGGTGAGACAAGTTTG
TCCCACAGGGCTGTCTGAGCAGATAAGATTAAGGGCTGGGTCTGTGCTCAATTAACCTCTGTGGG
CACGGGGGCTGGGAAGAGCAAAGTCAGCGGTGCCTACAGTCAGCACCATGCTGGGCCTGCCGTGG
AAGGGAGGTCTGTCTGGGCGCTGCTGCTGCTTCTCTTAGGCTCCCAGATCCTGCTGATCTATGC
CTGGCATTTCACGAGCAAAGGGACTGTGATGAACACAATGTCATGGCTCGTTACCTCCCTGCCA
CAGTGGAGTTTGTGTCCACACATTCAACCAACAGAGCAAGGACTACTATGCCTACAGACTGGGG
CACATCTTGAATTCCTGGAAGGAGCAGGTGGAGTCCAAGACTGTATTCTCAATGGAGCTACTGCT
GGGGAGAACTAGGTGTGGGAAATTTGAAGACGACATTGACAACTGCCATTTCCAAGAAAGCACAG
AGCTGAACAATACTTTCACCTGCTTCTTCACCATCAGCACCAGGCCCTGGATGACTCAGTTCAGC
CTCCTGAACAAGACCTGCTTGGAGGGATTCCACTTGAGTGAAACCCACTCACAGGCTTGTCCATGT
GCTGCTCCACATTCCGTGGACATCAGCACTACTCTCCTGAGGACTCTTCAGTGGCTGAGCAGCT
TTGGACTTGTTTGTATCTATTTTGCATGTGTTTGAGATCTCAGATCAGTGTTTTAGAAAATCC
ACACATCTTGAGCCTAATCATGTAGTGTAGATCATTAAACATCAGCATTTTAAGAAAAAAAAAA
AA

FIGURE 144

MLGLPWKGGLSWALLLLLLGSQILLIYAWHFHEQRDCDEHNVMARYLPATVEFAVHTFNQQSKDY
YAYRLGHILNSWKEQVESKTVFSMELLGRTRCGKFEDDIDNCHFQESTELNNTFTCFFTISTRP
WMTQFSLNKTCLGFFH

Important features of the protein:

Signal peptide:

amino acids 1-25

N-glycosylation sites.

amino acids 117-121, 139-143

N-myristoylation site.

amino acids 9-15

FIGURE 145

CTGTGCAGCTCGAGGCTCCAGAGGCACACTCCAGAGAGAGCCAAGGTTCCTGACGCGATGAGGAAG
CACCTGAGCTGGTGGTGGCTGGCCACTGTCTGCATGCTGCTCTTCAGCCACCTCTCTGCGGTCCA
GACGAGGGGCATCAAGCACAGAATCAAGTGGAAACCGGAAGGCCCTGCCCAGCACTGCCCAGATCA
CTGAGGCCCAGGTGGCTGAGAACCGCCCGGAGCCTTCATCAAGCAAGGCCGCAAGCTCGACATT
GACTTCGGAGCCGAGGGCAACAGGTACTACGAGGCCAACTACTGGCAGTTCCCCGATGGCATCCA
CTACAACGGCTGCTCTGAGGCTAATGTGACCAAGGAGGCATTTGTCACCGGCTGCATCAATGCCA
CCCAGGCGGGCAACCAGGGGGAGTTCCAGAAGCCAGACAACAAGCTCCACCAGCAGGTGCTCTGG
CGGCTGGTCCAGGAGCTCTGCTCCCTCAAGCATTGCGAGTTTTGGTTGGAGAGGGGCGCAGGACT
TCGGGTCACCATGCACCAGCCAGTGCTCCTCTGCCTTCTGGCTTTGATCTGGCTCATGGTGAAAT
AAGCTTGCCAGGAGGCTGGCAGTACAGAGCGCAGCAGCGAGCAAATCCTGGCAAGTGACCCAGCT
CTTCTCCCCCAAACCCACGCGTGTTCTGAAGGTGCCCAGGAGCGGCGATGCACTCGCACTGCAAA
TGCCGCTCCCACGTATGCGCCCTGGTATGTGCCTGCGTTCTGATAGATGGGGGACTGTGGCTTCT
CCGTCACTCCATTCTCAGCCCCTAGCAGAGCGTCTGGCACACTAGATTAGTAGTAAATGCTTGAT
GAGAAGAACACATCAGGCACTGCGCCACCTGCTTCACAGTACTTCCCAACAACCTCTTAGAGGTAG
GTGTATTCCCGTTTTACAGATAAGGAAACAGAGGCCAGAGAGCTGAAGTACTGCACCCAGCATC
ACCAGCTAGAAAGTGGCAGAGCCAGGATTCAACCCCTGGCTTGCTAACCCAGGTTTTCTGCTCT
GTCCAATTCCAGAGCTGTCTGGTGATCATTATGTCTCACAGGGACCCACATCCAAACATGTAT
CTCTAATGAAATTGTGAAAGCTCCATGTTTAGAAATAAATGAAAACACCTGA

FIGURE 146

MRKHLSWWLATVCMLLFSHLSAVQTRGIKHRIKWNRKALPSTAQITEAQAENRPGAFIKQGRK
LDIDFGAEGNRYYEANYWQFPDGIHYNGCSEANVTKEAFVTGCINATQAANQGEFQKPDNKLHQQ
VLWRLVQELCSLKHCFWLERGAGLRVTMHQPVLLCLLALIWMVK

Important features of the protein:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 157-171

N-glycosylation sites.

amino acids 98-102, 110-114

Tyrosine kinase phosphorylation site.

amino acids 76-83

N-myristoylation sites.

amino acids 71-77, 88-94, 93-99, 107-113, 154-160

Amidation site.

amino acids 62-66

FIGURE 147

GCCTTGGCCTCCCAAAGGGCTGGGATTATAGGCGTGACCACCATGTCTGGTCCAGAGTCTCATTT
CCTGATGATTTATAGACTCAAAGAAAACTGTTTCAGAAGCTCTCTTCTTTCTGGCCTCCTCT
CTGTCTTCTTTCCCTCTTTCTTCTTATTTTAATTAGTAGCATCTACTCAGAGTCATGCAAGCTGG
AAATCTTTCATTTTGCTTGTCAGTGGGGTAGGTCAGTCTTAGTTTTTATTTTTTGAAATTT
CAACTTTCAGATTCAGGGGGTACATGTGAAGGTTTGTTTTATGAGTATATTGCATGATGTGCTGAGG
TTTGGGGT

FIGURE 148

MFRSSLLFWPPLCLLSLFLLLISSIYSESKLEIFHFACQWGRSLSLSFYFLKFQLSDSGGTCE
GLFYEYIA

Important features of the protein:

Signal peptide:

amino acids 1-25

N-myristoylation site.

amino acids 62-68

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

FIGURE 149

GTCTCCGCGTCACAGGAACCTTCAGCACCCACAGGGCGGACAGCGCTCCCCTCTACCTGGAGACTTGAC
TCCCGCGCGCCCCAACCTTGCTTATCCCTTGACCGTCGAGTGTGAGAGATCCTGCAGCCGCCAGTCC
CGGCCCCCTCTCCGCCCCACACCCACCCTCCTGGCTCTTCTGTTTTACTCCTCCTTTTCATTGATA
ACAAAAGCTACAGCTCCAGGAGCCAGCGCCGGGCTGTGACCCAAGCCGAGCGTGAAGAATGGGGTT
CCTCGGGACCGGCACCTGGATTCTGGTGTAGTGTCTCCCGATTCAAGCTTTCCCCAACCTGGAGGAA
GCCAAGACAAATCTCTACATAATAGAGAATTAAGTGCAGAAAGACCTTTGAATGAACAGATTGCTGAA
GCAGAAGAAGACAAGATTAAAAAACATATCCTCCAGAAAACAAGCCAGGTCAGAGCAACTATTCTTT
TGTTGATAACTTGAACCTGCTAAAGGCAATAACAGAAAAGGAAAAAATTGAGAAAGAAAGACAATCTA
TAAGAAGCTCCCCACTTGATAATAAGTTGAATGTGGAAGATGTTGATTCAACCAAGAATCGAAAACTG
ATCGATGATTATGACTCTACTAAGAGTGGATTGGATCATAAATTTCAAGATGATCCAGATGGTCTTCA
TCACTAGACGGGACTCCTTTAACCGCTGAAGACATTGTCCATAAAATCGCTGCCAGGATTTATGAAG
AAAAATGACAGAGCCGTGTTTGACAAGATTGTTTCTAACTACTTAATCTCGGCCTTATCACAGAAAGC
CAAGCACATACACTGGAAGATGAAGTAGCAGAGGTTTACAAAAATTAATCTCAAAGGAAGCCAACAA
TTATGAGGAGGATCCCAATAAGCCCAAGCTGGACTGAGAATCAGGCTGGAAAAATACCAGAGAAAG
TGACTCCAATGGCAGCAATTCAAGATGGTCTTGCTAAGGGAGAAAACGATGAAACAGTATCTAACACA
TTAACCTTGACAAATGGCTTGAAAGGAGAACTAAAACCTACAGTGAAGACAACCTTTGAGGAACCTCCA
ATATTTCCCAAATTTCTATGCGCTACTGAAAAGTATTGATTGAGAAAAGAGCAAAAGAGAAAGAAA
CACTGATTACTATCATGAAAACACTGATTGACTTTGTGAAGATGATGGTGAATATGAACAATATCT
CCAGAAGAAGGTGTTTCTACCTTGAAAACCTGGATGAAATGATTGCTCTTCAGACCAAAAACAAGCT
AGAAAAAATGCTACTGACAATATAAGCAAGCTTTTCCAGCACCATCAGAGAAGAGTCATGAAGAAA
CAGACAGTACCAAGGAAGAAGCAGCTAAGATGGAAAAGGAATATGGAAGCTTGAAGGATTCCACAAAA
GATGATAACTCCAACCCAGGAGGAAAGACAGATGAACCCAAAGGAAAAACAGAAGCCTATTGGAAGC
CATCAGAAAAAATATTGAATGGTTGAAGAAACATGACAAAAGGGAATAAAGAAGATTATGACCTTT
CAAAGATGAGAGACTTCATCAATAAACAAGCTGATGCTTATGTGGAGAAAGGCATCCTTGACAAGGAA
GAAGCCGAGGCCATCAAGCGCATTATAGCAGCCTGTAAAAATGGCAAAGATCCAGGAGTCTTTCAA
CTGTTTCAGAAAACATAATATAGCTTAAACACTTCTAATTCTGTGATTAAAAATTTTTCACCAAGG
GTTATTAGAAAGTGCTGAATTTACAGTAGTTAACCTTTTACAAGTGGTTAAACATAGCTTTCTTCCC
GTAAAAACTATCTGAAAGTAAAGTTGTATGTAAGCTGAAAAAAAAAAAAAAAAAAAAA

FIGURE 150

MGFLGTGTWILVLVLP IQAFPKPGGSQDKSLHNRELSAERPLNEQIAEAEEDKIKKTYPPENKPG
QSNYSFVDNLNLLKAITEKEKIEKERQSIRSSPLDNKLNVEDVDSTKNRKLIDDYDSTKSGLDHK
FQDDPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVFDKIVSKLLNLGLITESQAHTLEDEVAE
VLQKLISKEANNYEEDPNKPTSWTENQAGK IPEKVT PMAAIQDGLAKGENDET VSNLT LTNGLE
RRTKTYSEDNFEELQYFPNFYALLKSIDSEKEAKEKETLITIMKTLIDFVKMMVKYGTISPEEGV
SYLENLDEMIALQTKNKLEKNATDNISKLFAPSEKSHEETDSTKEEAAKMEKEYGSLKDSTKDD
NSNPGGKTDEPKGKTEAYLEAIRKNIEWLKKHDKKGNKEDYDLSKMRDFINKQADAYVEKGILDK
EEAEAIKRIYSSL

N-glycosylation sites:

amino acids 68-71, 346-349, 350-353

Casein kinase II phosphorylation site:

amino acids 70-73, 82-85, 97-100, 125-128, 147-150, 188-191, 217-
220, 265-268, 289-292, 305-308, 320-323, 326-329, 362-365, 368-
341, 369-372, 382-385, 386-389, 387-390

N-myristoylation sites:

amino acids 143-148, 239-244

FIGURE 151

CGGCTCGAGGCTCCCGCCAGGAGAAAGGAACATTCTGAGGGGAGTCTACACCCTGTGGAGCTCAA
GATGGTCCTGAGTGGGGCGCTGTGCTTCCGAATGAAGGACTCGGCATTGAAGGTGCTTTATCTGC
ATAATAACCAGCTTCTAGCTGGAGGGCTGCATGCAGGGAAGGTCATTAAAGGTGAAGAGATCAGC
GTGGTCCCCAATCGGTGGCTGGATGCCAGCCTGTCCCCGTCATCCTGGGTGTCCAGGGTGGAAG
CCAGTGCTGTGCATGTGGGGTGGGGCAGGAGCCGACTCTAACACTAGAGCCAGTGAACATCATGG
AGCTCTATCTTGGTGCCAAGGAATCCAAGAGCTTCACCTTCTACCGGCGGGACATGGGGCTCACC
TCCAGCTTCGAGTCGGCTGCCCTACCGGGCTGGTTCCTGTGCACGGTGCCTGAAGCCGATCAGCC
TGTGAGACTCACCAGCTTCCCGAGAATGGTGGCTGGAATGCCCCATCACAGACTTCTACTTCC
AGCAGTGTGACTAGGGCAACGTGCCCCCAGAAGTCCCTGGGCAGAGCCAGCTCGGGTGAGGGGT
GAGTGGAGGAGACCCATGGCGGACAATCACTCTCTGCTCTCAGGACCCCCACGTCTGACTTAG
TGGGCACCTGACCACTTTGTCTTCTGGTTCACGTTTGGATAAATCTGAGATTGGAGCTCAGT
CCAGGTCCTCCCCCAGATGGTGGTCTACTGCTGTGGAACTTGTAAAAACCATGTGGGGTAAA
CTGGGAATAACATGAAAAGATTCTGTGGGGGTGGGGTGGGGGAGTGGTGGGAATCATTCCTGCT
TAATGGTAAGTACAAGTGTACCCTGAGCCCCGAGGCCAACCCATCCCCAGTTGAGCCTTATA
GGGTGAGTAGCTCTCCACATGAAGTCTGTCACTCACCCTGTGCAGGAGAGGGAGGTGGTCATA
GAGTCAGGGATCTATGGCCCTTGGCCAGCCCCACCCCTTCCCTTTAATCCTGCCACTGTCATA
TGCTACCTTTCTATCTCTTCCCTCATCATCTTGTGTGGGCATGAGGAGGTGGTGTATGTCAGAA
GAAATGGCTCGAGCTCAGAAGATAAAAGATAAGTAGGGTATGCTGATCCTCTTTAAAAACCCAA
GATACAATCAAAATCCCAGATGCTGGTCTCTATTCCCATGAAAAAGTGCTCATGACATATTGAGA
AGACCTACTTACAAAGTGGCATATATTGCAATTTATTTTAATTAAGATACCTATTTATATATT
TCTTTATAGAAAAAGTCTGGAAGAGTTTACTTCAATTGTAGCAATGTCAGGGTGGTGGCAGTAT
AGGTGATTTTTCTTTAATTCTGTTAATTTATCTGTATTTCCTAATTTTTCTACAATGAAGATGA
ATTCTTGTATAAAAAATAAGAAAAGAAATTAATCTTGAGGTAAGCAGAGCAGACATCATCTCTGA
TTGTCTCAGCCTCCACTTCCCCAGAGTAAATCAAATTGAATCGAGCTCTGCTGCTCTGGTTGG
TTGTAGTAGTGATCAGGAAACAGATCTCAGCAAAGCCACTGAGGAGGAGGCTGTGCTGAGTTTGT
GTGGCTGGAATCTCTGGGTAAGGAACCTTAAAGAACAAAAATCATCTGGTAATCTTTCTTAGAAG
GATCACAGCCCCCTGGGATTCGAAGGCATTGGATCCAGTCTCTAAGAAGGCTGCTGTACTGGTTGA
ATTGTGTCCCCCTCAAATTCACATCCTTCTGGAATCTCAGTCTGTGAGTTTATTTGGAGATAAG
GTCTCTGCAGATGTAGTTAGTTAAGACAAGGTCTGCTGGATGAAGGTAGACCTAAATTCAATAT
GACTGGTTTCCTTGTATGAAAAGGAGAGGACACAGAGACAGAGGAGACCGGGGAAGACTATGTA
AAGATGAAGGCAGAGATCGGAGTTTTCAGCCACAAGCTAAGAAACACCAAGGATTGTGGCAACC
ATCAGAAGCTTGAAGAGGCAAGAAGAATTCTTCCCTAGAGGCTTTAGAGGGATAACGGCTCTG
CTGAAACCTTAATCTCAGACTTCCAGCCTCCTGAACGAAGAAAGAATAAATTTGGCTGTTTTAA
GCCACCAAGGATAATTGGTTACAGCAGCTCTAGGAACTAATACAGCTGCTAAAATGATCCCTGT
CTCCTCGTGTTTACATTCTGTGTGTGCCCTCCCACATGTACCAAAGTTGTCTTTGTGACCAA
TAGAATATGGCAGAAGTGATGGCATGCCACTTCCAAGATTAGGTTATAAAAGACACTGCAGCTTC
TACTTGAGCCCTCTCTCTGTGCCACCCACCGCCCCAATCTATCTTGGCTCACTCGCTCTGGGGG
AAGCTAGCTGCCATGCTATGAGCAGGCCTATAAAGAGACTTACGTGGTAAAAAATGAAGTCTCCT
GCCACAGCCACATTAGTGAACCTAGAAGCAGAGACTCTGTGAGATAATCGATGTTTGTGTTTT
AAGTTGCTCAGTTTTTGGTCTAACTTGTATGCAGCAATAGATAAATAATATGCAGAGAAAGAG

Year	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100
1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100	

N-myristoylation sites.

Interleukin-1 signature.

Interleukin-1 proteins.

amino acids 8-29, 83-120, 95-134, 64-103

FIGURE 153

CTTCAGAACAGGTTCTCCCTTCCCCAGTCACCCAGTTGCTCGAGTTAGAATTGTCTGCAATGGCCGC
CCTGCAGAAATCTGTGAGCTCTTTCTTATGGGGACCTGGCCACCAGCTGCCTCCTTCTCTTG
CCCTCTTGGTACAGGGAGGAGCAGCTGCGCCCATCAGCTCCCACTGCAGGCTTGACAAGTCCAAC
TTCCAGCAGCCCTATATACCAACCGCACCTTCATGCTGGCTAAGGAGGCTAGCTTGGCTGATAA
CAACACAGACGTTCTGTCTCATTGGGGAGAACTGTTCCACGGAGTCAGTATGAGTGAGCGCTGCT
ATCTGATGAAGCAGGTGCTGAACCTTACCCTTGAAGAAGTGTGTTCCCTCAATCTGATAGGTT
CAGCCTTATATGCAGGAGGTGGTGGCCCTTCTGGCCAGGCTCAGCAACAGGCTAAGCACATGTCA
TATTGAAGGTGATGACCTGCATATCCAGAGGAATGTGCAAAGCTGAAGGACACAGTGAAAAAGC
TTGGAGAGAGTGGAGAGATCAAAGCAATTGGGAACTGGATTGCTGTTTATGTCTCTGAGAAAT
GCCTGCATTGACCAGAGCAAAGCTGAAAAATGAATAACTAACCCCTTCCCTGCTAGAAATAA
CAATTAGATGCCCCAAAGCGATTTTTTTTAAACCAAAGGAAGATGGGAAGCCAACTCCATCATG
ATGGGTGGATTCCAAATGAACCCCTGCGTTAGTTACAAAGGAAACCAATGCCACTTTGTTTATA
AGACCAGAAGGTAGACTTTCTAAGCATAGATATTTATTGATAACATTTTCATTGTAAGTGGTGTC
TATACACAGAAAACAATTTATTTTTTAAATAATTGTCTTTTCCATAAAAAAGATTACTTTCCAT
TCCTTTAGGGGAAAAAACCCCTAAATAGCTTCATGTTTCCATAATCAGTACTTTATATTTATAAA
TGTATTTATTATTATTATAAGACTGCATTTTATTTATATCATTTTATTAATATGGATTATTTAT
AGAAACATCATTCGATATTGCTACTTGAGTGTAAGGCTAATATTGATATTTATGACAATAATTAT
AGAGCTATAACATGTTTATTTGACCTCAATAAACACTTGGATATCCC

FIGURE 154

MAALQKSVSSFLMGTIATSCLLLLLALLVQGGAAAPISSHCRLDKSNFQQPYITNRTFMLAKEASL
ADNNTDVRLIGEKLFGVSMSERCYLMKQVLNFTLEEVLFPQSDRFQPYMQEVVPFLARLSNRLS
TCHIEGDDLHIQRNVQKLKDTVKKLGESGEIKAIGELDLLFMSLRNACI

Important features of the protein:

Signal peptide:

amino acids 1-33

N-glycosylation sites.

amino acids 54-58, 68-72, 97-101

N-myristoylation sites.

amino acids 14-20, 82-88

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 10-21

FIGURE 155

GGCTTGCTGAAAATAAAATCAGGACTCCTAACCTGCTCCAGTCAGCCTGCTTCCACGAGGCCTGT
CAGTCAGTGCCCGACTTGTGACTGAGTGTGAGTCAGTCAGCATGTACCAGGTCAGTGAGAGGGC
TGCCCTGAGGGCTGTGCTGAGAGGGAGAGGAGCAGAGATGCTGCTGAGGGTGGAGGGAGGCCAAGC
TGCCAGGTTTGGGGCTGGGGGCCAAGTGGAGTGAGAACTGGGATCCCAGGGGGAGGGTGAGAT
GAGGGAGCGACCCAGATTAGGTGAGGACAGTTCTCTCATTAGCCTTTTCTACAGGTGGTTGCAT
TCTTGGCAATGGTCATGGGAACCCACACCTACAGCCACTGGCCCAGCTGCTGCCCCAGCAAAGGG
CAGGACACCTCTGAGGAGCTGCTGAGGTGGAGCACTGTGCCTGTGCCTCCCCTAGAGCCTGCTAG
GCCCCAACCGCCACCCAGAGTCCTGTAGGGCCAGTGAAGATGGACCCCTCAACAGCAGGGCCATCT
CCCCCTGGAGATATGAGTTGGACAGAGACTTGAACCGGCTCCCCAGGACCTGTACCACGCCCCGT
TGCCCTGTGCCCCGACTGCGTCAGCCTACAGACAGGCTCCCACATGGACCCCCGGGGCAACTCGGA
GCTGCTCTACCACAACCAGACTGTCTTCTACAGCGGCCATGCCATGGCGAGAAGGGCACCCACA
AGGGCTACTGCCTGGAGCGCAGGCTGTACCGTGTTCCTTAGCTTGTGTGTGTGTGCGGCCCCGT
GTGATGGGCTAGCCGGACCTGCTGGAGGCTGGTCCCTTTTGGGAAACCTGGAGCCAGGTGTACA
ACCACTTGCCATGAAGGGCCAGGATGCCCAGATGCTTGGCCCCCTGTGAAGTGCTGTCTGGAGCAG
CAGGATCCCGGGACAGGATGGGGGGCTTTGGGGAAAACCTGCACTTCTGCACATTTTGAAAAGAG
CAGCTGCTGCTTAGGGCCGCCGGAAGCTGGTGTCTGTCATTTTCTCTCAGGAAAGGTTTCAAA
GTTCTGCCCCATTTCTGAGGCCACCACTCCTGTCTCTTCTCTTTTCCCATCCCCTGCTACCCTG
GCCCAGCACAGGCACTTTCTAGATATTTCCCCCTTGCTGGAGAAGAAAGAGCCCCCTGGTTTTATT
TGTTTGTTTACTCATCACTCAGTGAGCATCTACTTTGGGTGCATTCTAGTGTAGTTACTAGTCTT
TTGACATGGATGATTCTGAGGAGGAAGCTGTTATTGAATGTATAGAGATTTATCCAAATAAATAT
CTTTATTTAAAAATGAAAAA

FIGURE 156

MRERPRLGEDSSLISLFLQVVAFLAMVMGTHYSHWPSCCPKQDTSEELLRWSTVPVPPLEPA
RPNRHPESCRASEDGPLNSRAISPWRYELDRDLNRLPQDLYHARCLCPHCVSLQTGSHMDPRGNS
ELLYHNQTVFYRRPCHGEKGTHTKGYCLERRLYRVSLACVCVRPRVMG

Important features of the protein:

Signal peptide:

amino acids 1-32

N-glycosylation site.

amino acids 136-140

Tyrosine kinase phosphorylation site.

amino acids 127-135

N-myristoylation sites.

amino acids 44-50, 150-156

FIGURE 157

CCGGCGATGTCGCTCGTGTCTGCTAAGCCTGGCCGCGCTGTGCAGGAGCGCCGTACCCCGAGAGCC
GACCGTTCAATGTGGCTCTGAACTGGGCCATCTCCAGAGTGGATGCTACAACATGATCTAATCC
CCGGAGACTTGAGGGACCTCCGAGTAGAACCTGTTACAAC TAGTGTGCAACAGGGGACTATTCA
ATTTTGATGAATGTAAGCTGGGTACTCCGGGCAGATGCCAGCATCCGCTTGTGTAAGGCCACCA
GATTTGTGTGACGGGCAAAGCAACTTCCAGTCTACAGCTGTGTGAGGTGCAATTACACAGAGG
CCTTCCAGACTCAGACCAGACCCTCTGGTGGTAAATGGACATTTTCCTACATCGGCTTCCTGT
GAGCTGAACACAGTCTATTTTCATTGGGGCCATAATATTCCTAATGCAAATATGAATGAAGATGG
CCCTTCCATGTCTGTGAATTTACCTCACCAGGCTGCCTAGACCACATAATGAAATATAAAAAA
AGTGTGTCAAGGCCGGAAGCCTGTGGGATCCGAACATCACTGCTTGTGAAGAAGATGAGGAGACA
GTAGAAGTGAACCTCACAACCACTCCCTGGGAAACAGATACATGGCTCTTATCCAACACAGCAC
TATCATCGGGTTTTCTCAGGTGTTTGAGCCACACCAGAAGAAACAAACGCGAGCTTCAGTGGTGA
TTCCAGTGACTGGGGATAGTGAAGGTGCTACGGTGCAGCTGACTCCATATTTTCTACTTGTGGC
AGCGACTGCATCCGACATAAAGGAACAGTTGTGCTCTGCCACAAACAGGCGTCCCTTTCCCTCT
GGATAACAACAAAAGCAAGCCGGGAGGCTGGCTGCCTCTCCTCCTGCTGTCTGTGCTGGGCCA
CATGGGTGCTGGTGGCAGGGATCTATCTAATGTGGAGGCACGAAAGGATCAAGAAGACTTCCTTT
TCTACCACCACACTACTGCCCCCATTAAGGTTCTTGTGGTTTACCCATCTGAAATATGTTTCCA
TCACACAATTTGTTACTTCACTGAATTTCTTCAAACCATTCGAGAAGTGAGGTATCCTTGAAA
AGTGGCAGAAAAAGAAAATAGCAGAGATGGGTCCAGTGCAGTGGCTTGCCACTCAAAGAAGGCA
GCAGACAAAGTCGTCTTCCTTCTTTCCAATGACGTCAACAGTGTGTGCGATGGTACCTGTGGCAA
GAGCGAGGGCAGTCCCAGTGAGAACTCTCAAGACCTCTTCCCCCTTGCCCTTAAACCTTTCTGCA
GTGATCTAAGAAGCCAGATTCACTCTGCACAAATACGTGGTGGTCTACTTTAGAGAGATTGATACA
AAAGACGATTACAATGCTCTCAGTGTCTGCCCCAAGTACCACCTCATGAAGGATGCCACTGCTTT
CTGTGCAGAACTTCTCCATGTCAAGCAGCAGGTGTGAGCAGGAAAAAGATCACAAGCCTGCCACG
ATGGCTGCTGCTCCTTGTAG

FIGURE 158

MSLVLLSLAALCRSAVPREPTVQCGSETGPSPEWMLQHDLIPGDLRLDLRVEPVTTSVATGDYSILMNVSWV
LRADASIRLLKATKICVTGKSNFQSYSCVRCNYTEAFQTQTRPSGGKWTFSYIGFFVELNTVYFIGAHNIP
NANMNEDGPSMSVNFTSPGCLDHIMKYKKKCVKAGSLWDPNITACKKNEETVEVNFTTTPLGNRYMALIQH
STIIGFSQVFEPHQKKQTRASVVIPVTGDSEGATVQLTPYFPTCGSDCIRHKGTVVLCPTGVPFPLDNNK
SKPGGWLPLLLLSLLVATWVLVAGIYLMWRHERIKKTSFSTTTLLPPIKVLVVYPSEICFHHTICYFTEFL
QNHCRSEVILEKWQKKKIAEMGPVQWLATQKKAADKVVFLLSNDVNSVCDGTCGKSEGSPSENSQDLFPLA
FNLFCSDLRSQIHLHKYVVVYFREIDTKDDYNALSVCPKYHLMKDATAFCAELLHVKQQVSAGKRSQACHD
GCCSL

Important features of the protein:

Signal peptide:

amino acids 1-14

Transmembrane domain:

amino acids 290-309

N-glycosylation sites.

amino acids 67 - 71, 103 - 107, 156 - 160, 183 - 187, 197 - 201 and 283
- 287

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 228 - 232 and 319 - 323

Casein kinase II phosphorylation sites.

amino acids 178 - 182, 402 - 406, 414 - 418 and 453 - 457

N-myristoylation site.

amino acids 116-122

Amidation site.

amino acids 488-452

FIGURE 159

AGCCACCAGCGCAACATGACAGTGAAGACCCATGCATGGCCCAGCCATGGTCAAGTACTTGCTGCT
GTCGATATTGGGGCTTGCCTTTCTGAGTGAGGCGGCAGCTCGGAAAATCCCCAAAGTAGGACATA
CTTTTTTCCAAAAGCCTGAGAGTTGCCCGCCTGTGCCAGGAGGTAGTATGAAGCTTGACATTGGC
ATCATCAATGAAAACCAGCGCGTTTCCATGTCACGTAACATCGAGAGCCGCTCCACCTCCCCCTG
GAATTACACTGTCACTTGGGACCCCAACCGGTACCCCTCGGAAGTTGTACAGGCCCCAGTGTAGGA
ACTTGGGCTGCATCAATGCTCAAGGAAAGGAAGACATCTCCATGAATTCCGTTCCCATCCAGCAA
GAGACCCTGGTCGTCCGGAGGAAGCACCAAGGCTGCTCTGTTTCTTTCCAGTTGGAGAAGGTGCT
GGTGACTGTTGGCTGCACCTGCGTCACCCCTGTCATCCACCATGTGCAGTTAAGAGGTGCATATCC
ACTCAGCTGAAGAAG

FIGURE 16o

MTVKTLHGPA~~M~~VKYL~~L~~LSILGLAFLSEAAARKIPKVGHTFFQKPESC~~P~~PPVGGSMKLDIGIINEN
Q~~R~~VSM~~S~~RNIESRSTSPWNYT~~V~~TVWD~~P~~NRYPSEVVQAQCRNLGCINAQ~~G~~KEDISMNSVPIQ~~Q~~ETLVV
RRKHQGC~~S~~VSFQLEKVLVTVGCTC~~V~~TPVIHHVQ

Signal sequence:

amino acids 1-30

N-glycosylation site.

amino acids 83-87

N-myristoylation sites.

amino acids 106-111, 136-141

FIGURE 161

A C A C T G G C C A A A C A A A A C G A A A G C A C T C C G T G C T G G A A G T A G G A G G A G A G T C A G G A C T C C C A G G
 A C A G A G A G T G C A C A A A C T A C C C A G C A C A G C C C C C T C C G C C C C C T C T G G A G G C T G A A G A G G G A T T C
 C A G C C C C T G C C A C C C A C A G A C A C G G G C T G A C T G G G G T G T C T G C C C C C T T G G G G G G G G C A G C A C
 A G G G C C T C A G G C C T G G G T G C C A C C T G G C A C C T A G A A G A T G C C T G T G C C C T G G T T C T T G C T G T C C T
 T G G C A C T G G G C C G A A G C C C A G T G G T C C T T T C T C T G G A G A G G C T T G T G G G C C T C A G G A C G C T A C C
 C A C T G C T C T C C G G G C C T C T C C T G C C G C C T C T G G G A C A G T G A C A T A C T C T G C C T G C C T G G G G A C A T
 C G T G C C T G C T C C G G G C C C C G T G C T G G C G C C T A C G C A C C T G C A G A C A G A G C T G G T G C T G A G G T G C C
 A G A A G G A G A C C G A C T G T G A C C T C T G T C T G C G T G T G G C T G T C C A C T T G G C C G T G C A T G G G C A C T G G
 G A A G A G C C T G A A G A T G A G G A A A G T T T G G A G A G C A G C T G A C T C A G G G T G G A G G A G C C T A G G A A
 T G C C T C T C C A G G C C A A G T C G T G C T C T C C T T C C A G G C C T A C C C T A C T G C C C G C T G C G T C C T G C
 T G G A G G T G C A A G T G C C T G C T G C C C T T G T G C A G T T T G G T C A G T C T G T G G G C T C T G T G G T A T A T G A C
 T G C T T C G A G G C T G C C C T A G G G A G T G A G G T A C G A A T C T G G T C C T A T A C T C A G C C C A G G T A C G A G A A
 G G A A C T C A A C C A C A C A C A G C A G C T G C C T G C C C T G C C C T G G C T C A A C G T G T C A G C A G A T G G T G A C A
 A C G T G C A T C T G G T T C T G A A T G T C T C T G A G G A G C A G C A C T T C G G C C T C T C C C T G T A C T G G A A T C A G
 G T C C A G G G C C C C C A A A A C C C C G G T G G C A C A A A A C C T G A C T G G A C C G C A G A T C A T T A C C T T G A A
 C C A C A G A C C T G G T T C C C T G C C T C T G T A T T C A G G T G T G G C C T C T G G A A C C T G A C T C C G T T A G G A
 C G A A C A T C T G C C C C T T C A G G G A G G A C C C C G C G C A C A C C A G A A C C T C T G G C A A G C C G C C C G A C T G
 C G A C T G C T G A C C C T G C A G A G C T G G C T G C T G G A C G C A C C G T G C T C G T G C C C G C A G A A G C G G C A C T
 G T G C T G G C G G G C T C C G G G T G G G G A C C C C T G C C A G C C A C T G G T C C C A C C G C T T T C C T G G G A G A A C G
 T C A C T G T G G A C A A G G T T C T C G A G T T C C C A T T G C T G A A A G G C C A C C C T A A C C T C T G T G T T C A G G T G
 A A C A G C T C G G A G A A G C T G C A G C T G C A G G A G T G C T T G T G G G C T G A C T C C C T G G G G C C T C T C A A G A
 C G A T G T G C T A C T G T T G G A G A C A C G A G G C C C C A G G A C A A C A G A T C C C T C T G T G C C T T G G A A C C C A
 G T G G C T G T A C T T C A C T A C C C A G C A A A G C C T C C A C G A G G G C A G C T C G C C T T G G A G A G T A C T T A C T A
 C A A G A C C T G C A G T C A G G C C A G T G T C T G C A G C T A T G G G A C G A T G A C T T G G G A G C G C T A T G G G C C T G
 C C C C A T G G A C A A T A C A T C C A C A A G C G C T G G G C C C T C G T G T G G C T G G C C T G C C T A C T C T T T G C C G
 C T G C G C T T T C C C T C A T C C T C C T T C T C A A A A A G G A T C A C G C G A A A G G G T G G C T G A G G C T C T T G A A A
 C A G G A C G T C C G C T C G G G G C G G C C G C C A G G G C C G C G C G G C T C T G C T C C T A C T C A G C C G A T G A
 C T C G G G T T T C G A G C G C C T G G T G G G C G C C C T G G C G T C G G C C C T G T G C C A G C T G C C G C T G C C G C T G G
 C C G T A G A C C T G T G G A G C C G T C G T G A A C T G A G C G C G C A G G G G C C C G T G G C T T G G T T T C A C G C G C A G
 C G G C G C C A G A C C C T G C A G G A G G G C G G C G T G G T G G T C T T G C T C T C T C C C G G T G C G G T G G C G C T
 G T G C A G C G A G T G G C T A C A G G A T G G G G T G T C C G G G C C C G G G G C G C A G G C C C G C A C G A C G C C T T C C
 G C G C C T C G C T C A G C T G C G T G C T G C C C G A C T T C T T G C A G G G C C G G G C G C C C G C A G C T A C G T G G G G
 G C C T G C T T C G A C A G G C T G C T C C A C C C G G A C G C C G T A C C C G C C T T T T C G C A C C G T G C C C G T C T T
 C A C A C T G C C C T C C C A A C T G C C A G A C T T C C T G G G G G C C C T G C A G C A G C C T C G C G C C C C G C G T T C C G
 G G C G G C C C A A G A G A G A G C G G A G C A A G T G T C C C G G G C C C T T C A G C C A G C C C T G G A T A G C T A C T T C
 C A T C C C C G G G G A C T C C C G C G C C G G A C G C G G G T G G G A C C A G G G G C G G G A C C T G G G G C G G G G A
 C G G G A C T T A A A T A A A A G G C A G A C G C T G T T T T T C T A A A A A A

FIGURE 162

MPVPWFLLSLALGRSPVVLSELERLVGPQDATHCSPGLSCRLWDSILCLPGDIVPAPGPVLAPTHLQTELV
LRCQKETDCDLCLRVAVHLAVHGHWEPEDEEKFGGAADSGVEEPRNASLQAQVVLSFQAYPTARCVLLEV
QVPAALVQFGQSVGSVVYDCFEAALGSEVRIWSYTQPRYEKELNHTQQLPALPWLNVSAADGDNVHLVLNV
EEQHFGLSLYWNQVQGPPKPRWHKNTGPQIITLNHTDLVPCLCIQVWPLEPDSVRTNICPFREDPRAHQN
LWQAARLRLTLQSWLLDAPCSLPAEAALCWRAPGGDPCQPLVPPLSWENVTVDKVLEFPLLKGHPNLCVQ
VNSSEKLQEQECLWADSLGFLKDDVLLLETRGPQDNRSLEALEPSGCTSLPSKASTRAARLGEYLLQDLQS
GQCLQLWDDDLGALWACPMKYIHKRWALVWLACLLFAAALSLLLLKKDHAKGWLRLKQDVRSGAAARG
RAALLLYSADDSGFERLVGALASALCQLPLRVAVDLWSRRELSAQGPVAVFHAQRRQTLQEGGVVLLFSP
GAVALCSEWLQDGVSGPGAHGPHDAFRASLSCVLPDFLQGRAPGSYVGACFDRLLHPDAVPALFRTVPVFT
LPSQLPDFLALQOPRPRSRGLQERAEQVSRAALQPALDSYFHPPGTPAPGRGVGPGAGPGAGDGT

Signal sequence:

amino acids 1-20

Transmembrane domain.

amino acids 453-475

N-glycosylation sites.

amino acids 118-121, 186-189, 198-201, 211-214, 238-241, 248-251,
334-337, 357-360, 391-394

Glycosaminoglycan attachment site.

amino acids 583-586

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 552-555

N-myristoylation sites.

amino acids 107-112, 152-157, 319-324, 438-443, 516-521, 612-617,
692-697, 696-701, 700-705

FIGURE 163

GGGAGGGCTCTGTGCCAGCCCCGATGAGGACGCTGCTGACCATCTTGACTGTGGGATCCCTGGCT
 GCTCAGCCCCCTGAGGACCCCTCGGATCTGCTCCAGCACGTGAAATTCCAGTCCAGCAACTTTGA
 AAACATCCTGACGTGGGACAGCGGGCCAGAGGGCACCCAGACACGGTCTACAGCATCGAGTATA
 AGACGTACGGAGAGAGGGACTGGGTGGCAAAGAAGGGCTGTCAGCGGATCACCCGGAAGTCCTGC
 AACCTGACGGTGGAGACGGGCAACCTCACGGAGCTCTACTATGCCAGGGTCACCGCT
 GTCAGTGCGGGAGGCCGGTCAGCCACCAAGATGACTGACAGGTTGAGTCTCTGCAGCACACTAC
 CCTCAAGCCACCTGATGTGACCTGTATCTCCAAAGTGAGATCGATTGAGATGATTGTTTCATCCTA
 CCCCCACGCCAATCCGTGCAGGCGATGGCCACCGGCTAACCTGGAAGACATCTTCCATGACCTG
 TTCTACCACTTAGAGCTCCAGGTCAACCGCACCTACCAAATGCACCTTGAGGGGAAGCAGAGAGA
 ATATGAGTTCTTCGGCCTGACCCCTGACACAGAGTTCTTGGCACCATCATGATTGCGTTCCCA
 CCTGGGCCAAGGAGAGTGCCCCCTACATGTGCCGAGTGAAGACACTGCCAGACCGGACATGGACC
 TACTCCTTCTCCGGAGCCTTCTGTTCTCCATGGGCTTCTCGTCGAGTACTCTGTACCTGAG
 CTACAGATATGTCCCAAGCCGCTGCACCTCCCAACTCCCTGAACGTCAGCGAGTCTGACTT
 TCCAGCCGCTGCGCTTCATCCAGGAGCAGTCTGATCCCTGTCTTTGACCTCAGCGGCCCCAGC
 AGTCTGGCCCAGCCTGTCCAGTACTCCCAGATCAGGGTGTCTGGACCCAGGGAGCCCCGAGGAGC
 TCCACAGCGGCATAGCCTGTCCGAGATCACCTACTTAGGGCAGCCAGACATCTCCATCCTCCAGC
 CCTCCAACGTGCCACCTCCCCAGATCCTCTCCCCACTGTCTATGCCCCAAACGCTGCCCTGAG
 GTCGGGCCCCCATCTATGCACCTCAGGTGACCCCCGAAGCTCAATTCCCATCTACGCCCCACA
 GGCCATCTCTAAGGTCCAGCCTTCTCTATGCCCTCAAGCCACTCCGGACAGTGGCCTCCCT
 CCTATGGGGTATGCATGGAAGGTTCTGGCAAAGACTCCCCCACTGGGACACTTTCTAGTCTCTAAA
 CACCTTAGGCCTAAAGGTGAGCTTCAGAAAGAGCCACCAGCTGGAAGCTGCATGTTAGGTGGCCT
 TTCTCTGCAGGAGGTGACCTCCTTGGCTATGGAGGAATCCCAAGAAGCAAATCATTGCACCAGC
 CCCTGGGGATTTGCACAGACAGAACATCTGACCCAAATGTGCTACACAGTGGGGAGGAAGGGACA
 CCACAGTACCTAAAGGGCCAGCTCCCCCTCCTCTCCTCAGTCCAGATCGAGGGCCACCCCATGTC
 CCTCCCTTTGCAACCTCCTTCCGGTCCATGTTCCCCCTCGGACCAAGGTCCAAGTCCCTGGGGCC
 TGCTGGAGTCCCTTGTGTGTCCCAAGGATGAAGCCAAGAGCCAGCCCTGAGACCTCAGACCTG
 GAGCAGCCACAGAACTGGATTCTTTTTAGAGGCTGGCCCTGACTGTGCAGTGGGAGTCTCTG
 AGGGGAATGGGAAGGCTTGGTGCTTCTCCCTGTCCCTACCCAGTGTACATCCTTGGCTGTCA
 ATCCCATGCTGCCATGCCACACACTCTGCGATCTGGCCTCAGACGGGTGCCCTTGAGAGAAGC
 AGAGGGAGTGGCATGCAGGGCCCTGCCATGGGTGCGCTCCTCACCGGAACAAAGCAGCATGATA
 AGGACTGCAGCGGGGAGCTCTGGGGAGCAGCTTGTGTAGACAAGCGCGTGCTCGCTGAGCCCTG
 CAAGGCAGAAATGACAGTGCAAGGAGGAAATGCAGGGAACTCCCGAGGTCCAGAGCCCCACCTC
 CTAACACCATGGATTCAAAGTGCTCAGGGAATTTGCCTCTCCTTGCCCCATTCTTGCCAGTTTC
 ACAATCTAGCTCGACAGAGCATGAGGCCCCCTGCCTCTTCTGTGATTGTTCAAAGGTGGGAAGAGA
 GCCTGGAAAAGAACCAGGCCTGGAAAAGAACCAGAAGGAGGCTGGGCAGAACCAAGCAACCTGC
 ACTTCTGCCAAGGCCAGGGCCAGCAGGACGGCAGGACTCTAGGGAGGGGTGTGGCCTGCAGCTCA
 TTCCCAGCCAGGGCAACTGCCTGACGTTGCACGATTTAGCTTCATTCTCTGATAGAACAAAGC
 GAAATGCAGGTCCACCAGGGAGGGAGACACAAAGCCTTTTCTGCAGGCAGGAGTTTCAGACCCT
 ATCCTGAGAATGGGGTTTGAAGGAAGGTGAGGGCTGTGGCCCTTGACGGGTACAATAACACAC
 TGTACTGATGTCACAACTTTGCAAGCTCTGCCTTGGGTTGAGCCCATCTGGGCTCAAATTCAGC
 CTCACCACTCACAAGCTGTGTGACTTCAAACAAATGAAATCAGTGCCAGAACCTCGGTTTCTCTC
 ATCTGTAATGTGGGGATCATAACCTACCTCATGGAGTTTGGTGAAGATGAAATGAAGTCATG
 TCTTTAAAGTGCTTAATAGTGCTGGTACATGGGCAGTGCCCAATAAACGGTAGCTATTTAAAAA
 AAAAAAA

FIGURE 164

MRTLLTILTVGSLAAHAPEDPSDLLQHVKFQSSNFENILTWDSGPEGTPDTPVYSIEYKTYGERDW
VAKKGCQRITRKSCNLTVETGNLTelyyARVTAVSAGGRSATKMTDRFSSLQHTTLKPPDVTCTIS
KVRSIQMIVHPTPTPIRAGDGHRLTLEDIFHDLFYHLELQVNRTYQMHGGKQREYEFFGLTPDT
EFLGTIMICVPTWAKESAPYMCRVKTLDPDRTWTYSFSGAFLFSMGFLVAVLCYLSYRYVTKPPAP
PNSLNVQRVLTFFQPLRFIQEHVLIPVFDLSGPSSLAQPVQYSQIRVSGPREPAGAPQRHSLSEIT
YLGQPDISILQPSNVPPPQILSPLSYAPNAAPEVGPPSYAPQVTPEAQFFFYAPQAISKVQPSSY
APQATPDSWPPSYGVCMEGSGKDSPTGTLSSPKHLRPGQLQKEPPAGSCMLGGLSLQEVTSIAM
EESQEAKSLHQPLGICTDRTPDNVLHSGEETPQYLKGQLPLLSSVQIEGHPMSLPLQPPSGPC
SPSDQGPSPWGLLESVCPKDEAKSPAPETSDLEQPTELDSLFRGLALTIVQWES

Signal sequence.

amino acids 1-17

Transmembrane domain.

amino acids 233-250

N-glycosylation sites.

amino acids 80-83, 87-90, 172-175

N-myristoylation sites.

amino acids 11-16, 47-52, 102-107, 531-536, 565-570

FIGURE 165

TGGCCTACTGGAAAAAAAAAAAAAAAAAAAAAGTCACCCGGGGCCGCGGTGGCCACAACATGG
CTGCGGCGCCGGGGCTGCTCTTCTGGCTGTTTCGTGCTGGGGGCGCTCTGGTGGGTCCCGGGCCAG
TCGGATCTCAGCCACGGACGGCGTTTCTCGGACCTCAAAGTGTGCGGGGACGAAGAGTGCAGCAT
GTTAATGTACCGTGGGAAAGCTCTTGAAGACTTCACGGGCCCTGATTGTCGTTTTGTGAATTTTA
AAAAAGGTGACGATGTATATGTCTACTACAACTGGCAGGGGGATCCCTTGAACCTTTGGGCTGGA
AGTGTGAACACAGTTTTGGATATTTCCAAAAGATTTGATCAAGGTACTTCATAAATACACGGA
AGAAGAGCTACATATTCAGCAGATGAGACAGACTTTGTCTGCTTTGAAGGAGGAAGAGATGATT
TTAATAGTTATAATGTAGAAGAGCTTTTAGGATCTTTGGAAGTGGAGGACTCTGTACCTGAAGAG
TCGAAGAAAGCTGAAGAAGTTTCTCAGCACAGAGAGAAATCTCCTGAGGAGTCTCGGGGGCGTGA
ACTTGACCCTGTGCCGTAGCCCGAGGCATTGAGAGCTGATTGAGAGGATGGAGAAGGTGCTTTCT
CAGAGAGCACCGAGGGGCTGCAGGGACAGCCCTCAGCTCAGGAGAGCCACCTCACACCAGCGGT
CCTGCGGCTAACGCTCAGGGAGTGCAGTCTTCGTTGGACACTTTGAAGAAATTCGCACGATAA
ATTGAAAGTGCCGGGAAGCGAAAGCAGAACTGGCAATAGTTCTCCTGCCTCGGTGGAGCGGGAGA
AGACAGATGCTTACAAAGTCCTGAAAACAGAAATGAGTCAGAGAGGAAGTGGACAGTGCCTTATT
CATTACAGCAAAGGATTTTCGTTGGCATCAAATCTAAGTTTGTGTTTACAAAGATTGTTTTTAGTA
CTAAGCTGCCTTGGCAGTTTGCATTTTTGAGCCAAACAAAATATATTATTTCCCTTCTAAGTA
AAAAAAAAAAAAAAAAAAAA

FIGURE 166

MAAAPGLLFWLFLVGLALWWVPGQSDLSHGRRFSDLKVCGDEECSSMLMYRGKALEDFTGPD CRFVN
FKKGDDVYVYKLAGGSLELWAGSVEHSFGYFPKDLIKVLHKYTEELHIPAETDFVCFEGGRD
DFNSYNVEELLGSLELEDSPVEESKKAEEVSOHREKSPEESRGRELDPVPEPEAFRAEDGEDGA
FSESTGLQGQPSAQESHPTSGPAANAQGVQSSLDTFEEILHDKLKVPGSESRTGNSSPASVER
EKTDAYKVLKTEMSQRGSGQCVIHYSKGFRWHQNLSLFYKDCF

Important features of the protein:

Signal peptide:

amino acids 1-22

N-glycosylation site.

amino acids 294-298

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 30-34

Tyrosine kinase phosphorylation site.

amino acids 67-76

N-myristoylation sites.

amino acids 205-211, 225-231, 277-283

Amidation site.

amino acids 28-32

FIGURE 167

CCAGGACCAGGGCGCACC GGCTCAGCCTCTCACTTGTCTCAGAGGCCGGGGAAGAGAAGCAAAGCGC
AACGGTGTGGTCCAAGCCGGGGCTTCTGCTTCGCCCTCTAGGACATACACGGGACCCCTAACTTC
AGTCCCCCAAACGCGCACCTCGAAGTCTTGAACCTCCAGCCCCGCACATCCACGCGCGGCACAGG
CGCGGCAGGCGGCAGGTCCCGGCCGAAGGCGATGCGCGCAGGGGGTCGGGCAGCTGGGCTCGGGC
GGCGGGAGTAGGGCCCGGAGGGAGGAGGCTGCATATTCAGAGTCGCGGGCTGCGCCCTG
GGCAGAGGCCCGCTCGCTCCACGCAACACCTGCTGCTGCCACCGCGCCGCGATGAGCCGCGTGG
TCTCGCTGCTGCTGGGCGCGCGCTGCTGCTGCCGCCACGGAGCCTTCTGCCGCCGCTGGTCAGC
GGCCAAAAGGTGTGTTTTGCTGACTTCAAGCATCCCTGCTACAAAATGGCCTACTTCCATGAAC
GTCCAGCCGAGTGAGCTTTCAGGAGGCACGCTGGCTTGTGAGAGTGAGGGAGGAGTCCTCCTCA
GCCTTGAGAATGAAGCAGACAGAAAGTTAATAGAGAGCATGTTGCAAAACCTGACAAAACCCGGG
ACAGGGATTTCGTATGGTGATTTCTGGATAGGGCTTTGGAGGAATGGAGATGGGCAACATCTGG
TGCTGCCCAGATCTCTACAGTGGTCTGATGGAAGCAATCCCAGTACCGAAACTGGTACACAG
ATGAACCTTCCCGGGAAGTGAAAAGTGTGTTGTGATGTATCACCACCAACTGCCAATCCTGGC
CTTGGGGGTCCCTACCTTTACAGTGAATGATGACAGGTGAACATGAAGCACAATTATATTTG
CAAGTATGAACCAGAGATTAATCCAACAGCCCTGTAGAAAAGCCTTATCTTACAAATCAACCAG
GAGACCCCATCAGAATGTGGTGTGTTACTGAAGCAGGTATAATTCCCAATCTAATTTATGTTGTT
ATACCAACAATACCCCTGCTCTTACTGATACTGGTGTGCTTTTGAACCTGTTGTTTCCAGATGCT
GCATAAAAGTAAAGGAAGAACAAAACCTAGTCCAAACAGTCTACACTGTGGATTTCAAAGAGTA
CCAGAAAAGAAAGTGGCATGGAAGTATAAATACTCATTGACTTGGTTCCAGAATTTGTAAATCT
GGATCTGTATAAGGAATGGCATCAGAACAATAGCTTGAATGGCTTGAATCACAAGGATCTGC
AAGATGAACCTGTAAGCTCCCCCTTGAGGCAAAATATTAAGTAATTTTTATATGTCTATTATTTCA
TTTTAAAGATATGCTGTGCTAATAATGGAGTGAGACATGCTTATTTTGTAAAGGATGACCCAA
ACTTCAAACCTTCAAGCAAATGAAATGGACAATGCAGATAAAGTTGTTATCAACACGTCGGGAGTA
TGTGTGTTAGAAGCAATTCCTTTATTTCTTTACCTTTTCATAAGTTGTTATCTAGTCAATGTAA
TGTATATTTGATTGAAATTTACAGTGTGCAAAAGTATTTTACCTTTGCATAAGTGTGTTGATAAAA
ATGAACCTGTTCTAATATTTATTTTATGGCATCTCATTTTTCAATACATGCTCTTTTGATTAAAG
AACTTATTACTGTTGTCAACTGAATTCACACACACACAAATATAGTACCATAGAAAAAGTTTGT
TTTCTCGAAATAATTCATCTTTACGCTTCTCTGCTTTTGGTCAATGTCTAGGAAATCTCTCAGA
AATAAGAAGCTATTTTCAATTAAGTGTGATATAAACCTCCTCAAACATTTTACTTAGAGGCAAGGAT
TGTCTAATTTCAATTTGTGCAAGACATGTGCCTTATAATTATTTTAGCTTAAATTAACAGATT
TTGTAATAATGTAACCTTTGTTAATAGGTGCATAAACACTAATGCAGTCAATTTGAACAAAAGAG
TGACATACACAATATAAATCATATGTCTTCACACGTTGCCATATAATGAGAAGCAGCTCTCTGA
GGGTTCTGAAATCAATGTGGTCCCTCTCTTGCCCACTAAACAAAGATGGTTGTTGCGGGTTTGGG
ATTGACACTGGAGGCAGATAGTTGCAAAGTTAGTCTAAGTTTCCCTAGCTGTATTTAGCCTCTG
ACTATATTAGTATACAAAGAGGTGATGTGGTTGAGACCAGGTGAATAGTCACTATCAGTGTGGAG
ACAAGCACAGCACACAGACATTTTAGGAAGGAAAGGAACACGAAATCGTGTGAAATGGGTGG
AACCCATCAGTGATCGCATATTCATTGATGAGGGTTTGTGAGATAGAAAATGGTGGCTCCTTT
CTGTCTTATCTCCTAGTTTCTTCAATGCTTACGCTTGTCTTCTCAAGAGAAAGTTGTAACCTCT
CTGGTCTTCATATGTCCCTGTGCTCCTTTTAAACCAAATAAGAGTTCTTGTCTTCTGGGGGAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 168

MSRVVSLLLGAALLCGHGAFCRVVSQGKVCFADFKHPCYKMAYFHELSSRVSFQEARLACESE
GGVLLSLENEAEQKLIESMLQNLTKPGTGISDGDGFWIGLWRNGDGQTSGACPDLYQWSDGSNSQ
YRNWYTDEPSCGSEKCVVMYHQPTANPGLGGPYLYQWDDRCNMKHNYICKYEPEINPTAPVEK
PYLTNQPGDTHQNVVVTEAGIIPNLIYVVIPTIPLLLLLILVAFGTCCFQMLHKSCKGRTKTSPNQ
STLWISKSTRKESGMEV

Important features of the protein:

Signal peptide:

amino acids 1-21

Transmembrane domain:

amino acids 214-235

N-glycosylation sites.

amino acids 86-89, 255-258

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 266-269

N-myristoylation sites.

amino acids 27-32, 66-71, 91-96, 93-98, 102-107, 109-114, 140-145, 212-217